

HUMAN CDNAS AND PROTEINS AND USES THEREOF**RELATED APPLICATIONS**

[0001] The present application claims priority from U.S. Provisional Application Serial No. 60/311,305, filed August 10, 2001; U.S. Provisional Application Serial No. 60/314,734, filed August 5 24, 2001; U.S. Provisional Application Serial No. 60/318,204, filed September 7, 2001; and U.S. Provisional Application Serial No. 60/326,470, filed October 1, 2001 and entitled "Human cDNAs and Proteins and Uses Thereof," the disclosures of which are incorporated herein by reference in their entireties.

FIELD OF THE INVENTION

10 [0002] The present invention is directed to GENSET polynucleotides and polypeptides, and fragments, derivatives, and variants thereof. The present invention also relates to recombinant vectors including the polynucleotides of the present invention, particularly recombinant vectors comprising a GENSET gene regulatory region or a sequence encoding a GENSET polypeptide, and to host cells containing the polynucleotides of the invention. The invention further relates to antibodies that 15 specifically bind to the polypeptides of the invention and to methods for producing such antibodies and fragments thereof. The invention also provides for methods of detecting the presence of the polynucleotides and polypeptides of the present invention in a sample, methods of diagnosis and screening of abnormal GENSET polypeptide expression and/or biological activity, methods of screening compounds for their ability to modulate the activity or expression of the GENSET 20 polypeptides, and uses of such compounds.

BACKGROUND OF THE INVENTION

[0003] cDNAs encoding secreted proteins or fragments thereof represent a particularly valuable source of therapeutic agents. Thus, there is a need for the identification and characterization of secreted proteins and the nucleic acids encoding them.

25 [0004] In addition to being therapeutically useful themselves, secretory proteins include short peptides, called signal peptides, at their amino termini which direct their secretion. These signal peptides are encoded by the signal sequences located at the 5' ends of the coding sequences of genes encoding secreted proteins. Because these signal peptides will direct the extracellular secretion of any protein to which they are operably linked, the signal sequences may be exploited to direct the efficient 30 secretion of any protein by operably linking the signal sequences to a gene encoding the protein for which secretion is desired. In addition, fragments of the signal peptides called membrane-

translocating sequences may also be used to direct the intracellular import of a peptide or protein of interest. This may prove beneficial in gene therapy strategies in which it is desired to deliver a particular gene product to cells other than the cells in which it is produced. Signal sequences encoding signal peptides also find application in simplifying protein purification techniques. In such 5 applications, the extracellular secretion of the desired protein greatly facilitates purification by reducing the number of undesired proteins from which the desired protein must be selected. Thus, there exists a need to identify and characterize the 5' fragments of the genes for secretory proteins which encode signal peptides.

[0005] Sequences coding for secreted proteins may also find application as therapeutics or 10 diagnostics. In particular, such sequences may be used to determine whether an individual is likely to express a detectable phenotype, such as a disease, as a consequence of a mutation in the coding sequence for a secreted protein. In instances where the individual is at risk of suffering from a disease or other undesirable phenotype as a result of a mutation in such a coding sequence, the undesirable phenotype may be corrected by introducing a normal coding sequence using gene therapy. 15 Alternatively, if the undesirable phenotype results from overexpression of the protein encoded by the coding sequence, expression of the protein may be reduced using antisense or triple helix based strategies.

[0006] The secreted human polypeptides encoded by the coding sequences may also be used as therapeutics by administering them directly to an individual having a condition, such as a disease, 20 resulting from a mutation in the sequence encoding the polypeptide. In such an instance, the condition can be cured or ameliorated by administering the polypeptide to the individual.

[0007] In addition, the secreted human polypeptides or fragments thereof may be used to generate antibodies useful in determining the tissue type or species of origin of a biological sample. The antibodies may also be used to determine the cellular localization of the secreted human 25 polypeptides or the cellular localization of polypeptides which have been fused to the human polypeptides. In addition, the antibodies may also be used in immunoaffinity chromatography techniques to isolate, purify, or enrich the human polypeptide or a target polypeptide which has been fused to the human polypeptide.

SUMMARY OF THE INVENTION

30 [0008] The present invention provides a purified or isolated polynucleotide comprising, consisting of, or consisting essentially of a nucleotide sequence selected from the group consisting of: (a) the sequences of the **odd** SEQ ID NOS:1-23; (b) the sequences of clone inserts of the deposited

clone pool; (c) the coding sequences of the odd SEQ ID NOs:1-23; (d) the coding sequences of the clone inserts of the deposited clone pool; (e) the sequences encoding one of the polypeptides of the **even** SEQ ID NOs:2-24; (f) the sequences encoding one of the polypeptides encoded by the clone inserts of the deposited clone pool; (g) the genomic sequences coding for the GENSET polypeptides; 5 (h) the 5' transcriptional regulatory regions of GENSET genes; (i) the 3' transcriptional regulatory regions of GENSET genes; (j) the polynucleotides comprising the nucleotide sequence of any combination of (g)-(i); (k) the variant polynucleotides of any of the polynucleotides of (a)-(j); (l) the polynucleotides comprising a nucleotide sequence of (a)-(k), wherein the polynucleotide is single stranded, double stranded, or a portion is single stranded and a portion is double stranded; (m) the 10 polynucleotides comprising a nucleotide sequence complementary to any of the single stranded polynucleotides of (l). The invention further provides for fragments of the nucleic acids and polypeptides of (a)-(m) described above.

[0009] Further embodiments of the invention include purified or isolated polynucleotides that comprise, consist of, or consist essentially of a nucleotide sequence at least 70% identical, more 15 preferably at least 75%, and even more preferably at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identical, to any of the nucleotide sequences in (a)-(m) above, e.g. over a region of contiguous nucleotides at least about any one integer between 10 and the last integer representing the last integer representing the last nucleotide of a specified sequence of the sequence listing, or a polynucleotide which hybridizes under stringent hybridization conditions to a polynucleotide of the present invention 20 including (a) through (m) above.

[0010] The present invention also relates to recombinant vectors, which include the purified or isolated polynucleotides of the present invention, and to host cells recombinant for the polynucleotides of the present invention, as well as to methods of making such vectors and host cells. The present invention further relates to the use of these recombinant vectors and recombinant host 25 cells in the production of GENSET polypeptides. The present invention further relates to a polynucleotide of the present invention operably linked to a regulatory sequence including promoters, enhancers, etc.

[0011] The invention further provides a purified or isolated polypeptide comprising, consisting of, or consisting essentially of an amino acid sequence selected from the group consisting 30 of: (a) the full length polypeptides of **even** SEQ ID NOs:2-24; (b) the full length polypeptides encoded by the clone inserts of the deposited clone pool; (c) the epitope-bearing fragments of the polypeptides of even SEQ ID NOs:2-24; (d) the epitope-bearing fragments of the polypeptides encoded by the clone inserts contained in the deposited clone pool; (e) the domains of the polypeptides

of even SEQ ID NOs:2-24; (f) the domains of the polypeptides encoded by the clone inserts contained in the deposited clone pool; (g) the signal peptides of the polypeptides of even SEQ ID NOs:2-24 or encoded by the human cDNAs of the deposited clone pool; (h) the mature polypeptides of even SEQ ID NOs:2-24 or encoded by the human cDNAs of the deposited clone pool; and (i) the allelic variant 5 polypeptides of any of the polypeptides of (a)-(h). The invention further provides for fragments of the polypeptides of (a)-(i) above, such as those having biological activity or comprising biologically functional domain(s).

[0012] The present invention further includes polypeptides with an amino acid sequence with at least 70% similarity, and more preferably at least 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 10 99% similarity to those polypeptides described in (a)-(i), or fragments thereof, as well as polypeptides having an amino acid sequence at least 70% identical, more preferably at least 75% identical, and still more preferably 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identical to those polypeptides described in (a)-(i), or fragments thereof, e.g. over a region of amino acids at least any one integer between 6 and the last integer representing the last amino acid of a specified polypeptide sequence of 15 the sequence listing. The invention further relates to methods of making the polypeptides of the present invention.

[0013] The present invention further relates to transgenic plants or animals, wherein said transgenic plant or animal is transgenic for a polynucleotide of the present invention and expresses a polypeptide of the present invention.

20 [0014] The invention further relates to antibodies that specifically bind to GENSET polypeptides of the present invention and fragments thereof as well as to methods for producing such antibodies and fragments thereof.

[0015] The invention also provides kits, uses and methods for detecting GENSET gene expression and/or biological activity in a biological sample. One such method involves assaying for 25 the expression of a GENSET polynucleotide in a biological sample using the polymerase chain reaction (PCR) to amplify and detect GENSET polynucleotides or Southern and Northern blot hybridization to detect GENSET genomic DNA, cDNA or mRNA. Alternatively, a method of detecting GENSET gene expression in a test sample can be accomplished using a compound which binds to a GENSET polypeptide of the present invention or a portion of a GENSET polypeptide.

30 [0016] The present invention also relates to diagnostic methods and uses of GENSET polynucleotides and polypeptides for identifying individuals or non-human animals having elevated or reduced levels of GENSET gene products, which individuals are likely to benefit from therapies to

suppress or enhance GENSET gene expression, respectively, and to methods of identifying individuals or non-human animals at increased risk for developing, or at present having, certain diseases/disorders associated with GENSET polypeptide expression or biological activity.

[0017] The present invention also relates to kits, uses and methods of screening compounds 5 for their ability to modulate (e.g. increase or inhibit) the activity or expression of GENSET polypeptides including compounds that interact with GENSET gene regulatory sequences and compounds that interact directly or indirectly with a GENSET polypeptide. Uses of such compounds are also within the scope of the present invention.

[0018] The present invention also relates to pharmaceutical or physiologically acceptable 10 compositions comprising, an active agent, the polypeptides, polynucleotides or antibodies of the present invention, as well as, typically, a physiologically acceptable carrier.

[0019] The present invention also relates to computer systems containing cDNA codes and polypeptide codes of sequences of the invention and to computer-related methods of comparing 15 sequences, identifying homology or features using GENSET polypeptides or GENSET polynucleotide sequences of the invention.

[0020] In another aspect, the present invention provides an isolated polynucleotide, the polynucleotide comprising a nucleic acid sequence encoding a polypeptide of the present invention including the polypeptide of (a) through (i) above.

[0021] In another aspect, the present invention provides a non-human transgenic animal 20 comprising the host cell.

[0022] In another aspect, the present invention provides a method of making a GENSET polypeptide, the method comprising a) providing a population of host cells comprising a herein-described polynucleotide and b) culturing the population of host cells under conditions conducive to the production of the polypeptide within said host cells.

25 [0023] In one embodiment, the method further comprises purifying the polypeptide from the population of host cells.

[0024] In another aspect, the present invention provides a method of making a GENSET polypeptide, the method comprising a) providing a population of cells comprising a polynucleotide encoding a herein-described polypeptide; b) culturing the population of cells under conditions 30 conducive to the production of the polypeptide within the cells; and c) purifying the polypeptide from the population of cells.

[0025] In another aspect, the present invention provides a biologically active polypeptide encoded by any of the herein-described polynucleotides.

[0026] In one embodiment, the polypeptide is selectively recognized by an antibody raised against an antigenic polypeptide, or an antigenic fragment thereof, the antigenic polypeptide comprising any one of the sequences shown as even SEQ ID NOS:2-24 or any one of the sequences of polypeptides encoded by the human cDNAs of the deposited clone pool.

[0027] In another aspect, the present invention provides an antibody that specifically binds to any of the herein-described polypeptides and methods of binding antibody to said polypeptide.

[0028] In another aspect, the present invention provides a method of determining whether a GENSET gene is expressed within a mammal, the method comprising the steps of: a) providing a biological sample from said mammal; b) contacting said biological sample with either of: (i) a polynucleotide that hybridizes under stringent conditions to any of the herein-described polynucleotides; or (ii) a polypeptide that specifically binds to any of the herein-described polypeptides; and c) detecting the presence or absence of hybridization between the polynucleotide and an RNA species within the sample, or the presence or absence of binding of the polypeptide to a protein within the sample; wherein a detection of the hybridization or of the binding indicates that the GENSET gene is expressed within the mammal.

[0029] In one embodiment, the polynucleotide is a primer, and the hybridization is detected by detecting the presence of an amplification product comprising the sequence of the primer. In another embodiment, the polypeptide is an antibody.

[0030] In another aspect, the present invention provides a method of determining whether a mammal has an elevated or reduced level of GENSET gene expression, the method comprising the steps of: a) providing a biological sample from the mammal; and b) comparing the amount of any of the herein-described polypeptides, or of an RNA species encoding the polypeptide, within the biological sample with a level detected in or expected from a control sample; wherein an increased amount of the polypeptide or the RNA species within the biological sample compared to the level detected in or expected from the control sample indicates that the mammal has an elevated level of the GENSET gene expression, and wherein a decreased amount of the polypeptide or the RNA species within the biological sample compared to the level detected in or expected from the control sample indicates that the mammal has a reduced level of the GENSET gene expression.

[0031] In another aspect, the present invention provides a method of identifying a candidate modulator of a GENSET polypeptide, the method comprising: a) contacting any of the herein-

described polypeptides with a test compound; and b) determining whether the compound specifically binds to the polypeptide; wherein a detection that the compound specifically binds to the polypeptide indicates or inhibits or activates of a specified biological activity that the compound is a candidate modulator of the GENSET polypeptide.

5

BRIEF DESCRIPTION OF TABLES

[0032] **Table I** provides the Applicants' internal designation number (Clone ID_Clone Name) which corresponds to each sequence identification number (SEQ ID NO) of the Sequence Listing, and indicates whether the sequence is a nucleic acid sequence (DNA) or a polypeptide sequence (PRT).

Further provided is information regarding the name of the corresponding nucleic acid or polypeptide 10 sequence, and information regarding the deposit of biological material. It should be appreciated that biological materials have been deposited with reference to their corresponding Clone ID, Clone Name, or both Clone ID_Clone Name.

[0033] **Table II** provides the positions of the nucleotides of the corresponding SEQ ID NOs of the Sequence Listing which comprise the open reading frame (ORF), signal peptide, mature peptide, 15 polyadenylation signal, and the polyA tail of the polynucleotides of the invention.

[0034] **Table III** provides the positions of the amino acid of the corresponding SEQ ID NOs. of the Sequence Listing which comprise the positions of immunogenic epitopes of the polypeptides of the invention, which are useful in antibody generation.

BRIEF DESCRIPTION OF SEQUENCES

20 [0035] Sequences are presented in the accompanying Sequence Listing.

[0036] Odd SEQ ID NOs:1-23 are the nucleotide sequences of cDNAs, with open reading frames as indicated. When appropriate, the potential polyadenylation site and polyadenylation signal are also indicated.

[0037] Even SEQ ID NOs:2-24 are the amino acid sequences of proteins encoded by the 25 cDNAs of odd SEQ ID NOs:1-23.

[0038] In accordance with the regulations relating to Sequence Listings, the following codes have been used in the Sequence Listing to describes nucleotide sequences. The code "r" in the sequences indicates that the nucleotide may be a guanine or an adenine. The code "y" in the sequences indicates that the nucleotide may be a thymine or a cytosine. The code "m" in the sequences indicates 30 that the nucleotide may be an adenine or a cytosine. The code "k" in the sequences indicates that the nucleotide may be a guanine or a thymine. The code "s" in the sequences indicates that the nucleotide

may be a guanine or a cytosine. The code "w" in the sequences indicates that the nucleotide may be an adenine or a thymine. In addition, all instances of the symbol "n" in the nucleic acid sequences mean that the nucleotide can be adenine, guanine, cytosine or thymine.

[0039] In some instances, the polypeptide sequences in the Sequence Listing contain the symbol "Xaa." These "Xaa" symbols indicate either (1) a residue which cannot be identified because of nucleotide sequence ambiguity or (2) a stop codon in the determined sequence where applicants believe one should not exist (if the sequence were determined more accurately). In some instances, several possible identities of the unknown amino acids may be suggested by the genetic code.

[0040] In the case of secreted proteins, it should be noted that, in accordance with the regulations governing Sequence Listings, in the appended Sequence Listing the encoded protein (i.e. the protein containing the signal peptide and the mature protein or fragment thereof) extends from an amino acid residue having a negative number through a positively numbered amino acid residue. Thus, the first amino acid of the mature protein resulting from cleavage of the signal peptide is designated as amino acid number 1, and the first amino acid of the signal peptide is designated with the appropriate negative number.

[0041] In the case that a polynucleotide or polypeptide sequence described in the specification for SEQ ID NOs:1-24 is in conflict with the corresponding sequence provided in the Sequence listing, the sequences provided in the Sequence listing controls.

[0042] It should be appreciated that the polynucleotide and polypeptide sequences of SEQ ID NO:1-24 of the Sequence Listing are hereby incorporated by reference in their entireties.

DETAILED DESCRIPTION OF THE INVENTION AND PREFERRED EMBODIMENTS

Definitions.

[0043] Before describing the invention in greater detail, the following definitions are set forth to illustrate and define the meaning and scope of the terms used to describe the invention herein.

[0044] The term "GENSET gene," when used herein, encompasses genomic, mRNA and cDNA sequences encoding a GENSET polypeptide, including the 5' and 3' untranslated regions of said sequences.

[0045] The term "GENSET polypeptide biological activity" or "GENSET biological activity" is intended for polypeptides exhibiting any activity similar, but not necessarily identical, to an activity

of a GENSET polypeptide of the invention. The GENSET polypeptide biological activity of a given polypeptide may be assessed using any suitable biological assay, a number of which are known to those skilled in the art. In contrast, the term "biological activity" refers to any activity that any polypeptide may have.

5 [0046] The term "corresponding mRNA" refers to mRNA which was or can be a template for cDNA synthesis for producing a cDNA of the present invention.

[0047] The term "corresponding genomic DNA" refers to genomic DNA which encodes an mRNA of interest, e.g. corresponding to a cDNA of the invention, which genomic DNA includes the sequence of one of the strands of the mRNA, in which thymidine residues in the sequence of the 10 genomic DNA (or cDNA) are replaced by uracil residues in the mRNA.

[0048] The term "deposited clone pool" is used herein to refer to the pool of clones deposited with the ATCC or other depositary authority.

[0049] The term "heterologous", when used herein in reference to a polypeptide or polynucleotide, is intended to designate any polynucleotide or polypeptide other than a particular 15 GENSET polynucleotide or GENSET polypeptide of the invention, respectively.

[0050] "Providing" with respect to, e.g. a biological sample, population of cells, etc. indicates that the sample, population of cells, etc. is somehow used in a method or procedure. Significantly, "providing" a biological sample or population of cells does not require that the sample or cells are specifically isolated or obtained for the purposes of the invention, but can instead refer, for example, to 20 the use of a biological sample obtained by another individual, for another purpose.

[0051] An "amplification product" refers to a product of any amplification reaction, e.g. PCR, RT-PCR, LCR, etc.

[0052] A "modulator" of a protein or other compound refers to any agent that has a functional effect on the protein, including physical binding to the protein, alterations of the quantity or quality of 25 expression of the protein, altering any measurable or detectable activity, property, or behavior of the protein, or in any way interacts with the protein or compound.

[0053] "A test compound" can be any molecule that is evaluated for its ability to modulate a protein or other compound.

[0054] An antibody or other compound that specifically binds to a polypeptide or 30 polynucleotide of the invention is also said to "selectively recognize" the polypeptide or polynucleotide.

[0055] The term "isolated" with respect to a molecule requires that the molecule be removed from its original environment (e. g., the natural environment if it is naturally occurring). For example, a naturally-occurring polynucleotide or polypeptide present in a living animal is not isolated, but the same polynucleotide or polypeptide, separated from some or all of the coexisting materials in the 5 natural system, is isolated. Such polynucleotide could be part of a vector and/or such polynucleotide or polypeptide could be part of a composition, and still be isolated in that the vector or composition is not part of its natural environment. Specifically excluded from the definition of "isolated" are: naturally-occurring chromosomes (such as chromosome spreads), artificial chromosome libraries, genomic libraries, and cDNA libraries that exist either as an *in vitro* nucleic acid preparation or as a 10 transfected/transformed host cell preparation, wherein the host cells are either an *in vitro* heterogeneous preparation or plated as a heterogeneous population of single colonies. Also specifically excluded are the above libraries wherein a specified polynucleotide makes up less than 5% (may also be specified as 10%, 25%, 50%, or 75%) of the number of nucleic acid inserts in the vector molecules. Further specifically excluded are whole cell genomic DNA or whole cell RNA 15 preparations (including said whole cell preparations which are mechanically sheared or enzymatically digested). Further specifically excluded are the above whole cell preparations as either an *in vitro* preparation or as a heterogeneous mixture separated by electrophoresis (including blot transfers of the same) wherein the polynucleotide of the invention has not further been separated from the heterologous polynucleotides in the electrophoresis medium (e.g., further separating by excising a 20 single band from a heterogeneous band population in an agarose gel or nylon blot).

[0056] The term "purified" does not require absolute purity; rather, it is intended as a relative definition. Purification of starting material or natural material to at least one order of magnitude, preferably two or three orders, and more preferably four or five orders of magnitude is expressly contemplated. The term "purified" is further used herein to describe a polypeptide or polynucleotide 25 of the invention which has been separated from other compounds including, but not limited to, polypeptides or polynucleotides, carbohydrates, lipids, etc. The term "purified" may be used to specify the separation of monomeric polypeptides of the invention from oligomeric forms such as homo- or hetero- dimers, trimers, etc. The term "purified" may also be used to specify the separation of covalently closed (i.e. circular) polynucleotides from linear polynucleotides. A substantially pure 30 polypeptide or polynucleotide typically comprises about 50%, preferably 60 to 90% weight/weight of a polypeptide or polynucleotide sample, respectively, more usually about 95%, and preferably is over about 99% pure but, may be specified as any integer of percent between 50 and 100. Polypeptide and polynucleotide purity, or homogeneity, is indicated by a number of means well known in the art, such

as agarose or polyacrylamide gel electrophoresis of a sample, or using HPLC. As an alternative embodiment, purification of the polypeptides and polynucleotides of the present invention may be expressed as "at least" a percent purity relative to heterologous polypeptides and polynucleotides (DNA, RNA or both). As a preferred embodiment, the polypeptides and polynucleotides of the present invention
5 are at least; 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 96%, 96%, 98%, 99%, or 100% pure relative to heterologous polypeptides and polynucleotides, respectively. As a further preferred embodiment the polypeptides and polynucleotides have a purity ranging from any number, to the thousandth position, between 90% and 100% (e.g., a polypeptide or polynucleotide at least 99.995% pure)
10 relative to either heterologous polypeptides or polynucleotides, respectively, or as a weight/weight ratio relative to all compounds and molecules other than those existing in the carrier. Each number representing a percent purity, to the thousandth position, may be claimed as individual species of purity.

[0057] As used interchangeably herein, the terms "nucleic acid molecule(s)", "oligonucleotide(s)", and "polynucleotide(s)" include RNA or DNA (either single or double stranded,
15 coding, complementary or antisense), or RNA/DNA hybrid sequences of more than one nucleotide in either single chain or duplex form (although each of the above species may be particularly specified). The term "nucleotide" is used herein as an adjective to describe molecules comprising RNA, DNA, or RNA/DNA hybrid sequences of any length in single-stranded or duplex form. More precisely, the expression "nucleotide sequence" encompasses the nucleic material itself and is thus not restricted to
20 the sequence information (i.e. the succession of letters chosen among the four base letters) that biochemically characterizes a specific DNA or RNA molecule. The term "nucleotide" is also used herein as a noun to refer to individual nucleotides or varieties of nucleotides, meaning a molecule, or individual unit in a larger nucleic acid molecule, comprising a purine or pyrimidine, a ribose or deoxyribose sugar moiety, and a phosphate group, or phosphodiester linkage in the case of nucleotides
25 within an oligonucleotide or polynucleotide. The term "nucleotide" is also used herein to encompass "modified nucleotides" which comprise at least one modification such as (a) an alternative linking group, (b) an analogous form of purine, (c) an analogous form of pyrimidine, or (d) an analogous sugar (see, e.g., WO 95/04064). Preferred modifications of the present invention include, but are not limited to, 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4-acetylcytosine,
30 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine,

5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v) ybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid, 5 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, and 2,6-diaminopurine. The polynucleotide sequences of the invention may be prepared by any known method, including synthetic, recombinant, *ex vivo* generation, or a combination thereof, as well as utilizing any purification methods known in the art (for teachings regarding the preparation of modified oligos and nucleotides, see, e.g., U.S. Pat. Nos. 5,378,825; 5,386,023; 5,489,677; 5,602,240; and 5,610,289, U.S. Pat. Nos. 10 5,264,562 and 5,264,564, U.S. Pat. No. 5,223,618, U.S. Pat. No. 5,508,270, U.S. Pat. No. 4,469,863, U.S. Pat. Nos. 5,610,289 or 5,625,050, U.S. Pat. No. 5,256,775 or U.S. Pat. No. 5,366,878, PCT applications WO 94/17093 and WO 94/02499, U.S. Pat. No. 5,476,925, U.S. Pat. No. 5,023,243, U.S. Pat. Nos. 5,130,302 and 5,177,198).

[0058] The term “upstream” is used herein to refer to a location which is toward the 5' end of 15 the polynucleotide from a specific reference point.

[0059] The terms “complementary” or “complement thereof” are used herein to refer to the sequences of polynucleotides which is capable of forming Watson & Crick base pairing with another specified polynucleotide throughout the entirety of the complementary region. For the purpose of the present invention, a first polynucleotide is deemed to be complementary to a second polynucleotide 20 when each base in the first polynucleotide is paired with its complementary base. Complementary bases are, generally, A and T (or A and U), or C and G. “Complement” is used herein as a synonym from “complementary polynucleotide”, “complementary nucleic acid” and “complementary nucleotide sequence”. These terms are applied to pairs of polynucleotides based solely upon their sequences and not any particular set of conditions under which the two polynucleotides would actually bind. Unless 25 otherwise stated, all complementary polynucleotides are fully complementary on the whole length of the considered polynucleotide.

[0060] The terms “polypeptide” and “protein”, used interchangeably herein, refer to a polymer of amino acids without regard to the length of the polymer; thus, peptides, oligopeptides, and proteins are included within the definition of polypeptide. This term also does not specify or exclude 30 chemical or post-expression modifications of the polypeptides of the invention, although chemical or post-expression modifications of these polypeptides may be included or excluded as specific embodiments. Therefore, for example, modifications to polypeptides that include the covalent attachment of glycosyl groups, acetyl groups, phosphate groups, lipid groups and the like are expressly

encompassed by the term polypeptide. Further, polypeptides with these modifications may be specified as individual species to be included or excluded from the present invention. The natural or other chemical modifications, such as those listed in examples above can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl 5 termini, and may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched, for example, as a result of ubiquitination, and they may be cyclic, with or without branching. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of 10 flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of cysteine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA 15 mediated addition of amino acids to proteins such as arginylation, and ubiquitination (see, for instance Creighton, (1993), Posttranslational Covalent Modification of Proteins, W.H. Freeman and Company, New York B.C. Johnson, Ed., Academic Press, New York 1-12; Seifter, *et al.*, (1990) Meth Enzymol 182:626-646; Rattan *et al.*, (1992) Ann NY Acad Sci 663:48-62). Also included within the definition are polypeptides which contain one or more analogs of an amino acid (including, for example, non- 20 naturally occurring amino acids, amino acids which only occur naturally in an unrelated biological system, modified amino acids from mammalian systems, etc.), polypeptides with substituted linkages, as well as other modifications known in the art, both naturally occurring and non-naturally occurring.

[0061] As used herein, the terms "recombinant polynucleotide" and "polynucleotide construct" are used interchangeably to refer to linear or circular, purified or isolated polynucleotides 25 that have been artificially designed and which comprise at least two nucleotide sequences that are not found as contiguous nucleotide sequences in their initial natural environment. In particular, these terms mean that the polynucleotide or cDNA is adjacent to "backbone" nucleic acid to which it is not adjacent in its natural environment. Additionally, to be "enriched" the cDNAs will represent 5% or more of the number of nucleic acid inserts in a population of nucleic acid backbone molecules. 30 Backbone molecules according to the present invention include nucleic acids such as expression vectors, self-replicating nucleic acids, viruses, integrating nucleic acids, and other vectors or nucleic acids used to maintain or manipulate a nucleic acid insert of interest. Preferably, the enriched cDNAs represent 15% or more of the number of nucleic acid inserts in the population of recombinant

backbone molecules. More preferably, the enriched cDNAs represent 50% or more of the number of nucleic acid inserts in the population of recombinant backbone molecules. In a highly preferred embodiment, the enriched cDNAs represent 90% or more (including any number between 90 and 100%, to the thousandth position, e.g., 99.5%) of the number of nucleic acid inserts in the population 5 of recombinant backbone molecules.

[0062] The term "recombinant polypeptide" is used herein to refer to polypeptides that have been artificially designed and which comprise at least two polypeptide sequences that are not found as contiguous polypeptide sequences in their initial natural environment, or to refer to polypeptides which have been expressed from a recombinant polynucleotide.

10 **[0063]** As used herein, the term "operably linked" refers to a linkage of polynucleotide elements in a functional relationship. A sequence which is "operably linked" to a regulatory sequence such as a promoter means that said regulatory element is in the correct location and orientation in relation to the nucleic acid to control RNA polymerase initiation and expression of the nucleic acid of interest. For instance, a promoter or enhancer is operably linked to a coding sequence if it affects the 15 transcription of the coding sequence.

[0064] The term "domain" refers to an amino acid fragment with specific biological properties. This term encompasses all known structural and linear biological motifs. Examples of such motifs include but are not limited to leucine zippers, helix-turn-helix motifs, glycosylation sites, ubiquitination sites, alpha helices, and beta sheets, signal peptides which direct the secretion of 20 proteins, sites for post-translational modification, enzymatic active sites, substrate binding sites, and enzymatic cleavage sites.

[0065] Although each of these terms has a distinct meaning, the terms "comprising", "consisting of" and "consisting essentially of" may be interchanged for one another throughout the instant application. The term "having" has the same meaning as "comprising" and may be replaced 25 with either the term "consisting of" or "consisting essentially of".

[0066] Unless otherwise specified in the application, nucleotides and amino acids of polynucleotides and polypeptides, respectively, of the present invention are contiguous and not interrupted by heterologous sequences.

[0067] As used herein, the term "tumor" refers to an abnormal mass or population of cells that 30 result from excessive cell division, whether malignant or benign, and all pre-cancerous and cancerous cells and tissues. A "tumor" is further defined as two or more physically associated neoplastic cells. The term "transformed cells," "malignant cells" or "cancer" are interchangeable and refer to cells that

have undergone malignant transformation, and also includes lymphocytes that have undergone blast transformation. Transformed cells have a greater ability to cause tumors when injected into animals, can typically proliferate without requiring adhesion to a substratum, and also lack contact inhibition. The term "cancer" or "neoplastic disease" encompasses any type of cancer, in any tissue, and includes,
5 but is not limited to, carcinomas, lymphomas, blastomas, sarcomas, and leukemias.

[0068] The terms "inducing" or "inducing" with respect to a cellular process, e.g., apoptosis, refers to increasing or decreasing the number of cells that undergo the process, or the rate by which cells undergo the process, in a given cell population. Preferably the increase or decrease is at least 1.25, 1.5, 2, 5, 10, 50, 100, 500 or 1000 fold increase or decrease as compared to normal, untreated or
10 negative control cells.

[0069] A "therapeutically effective amount", in reference to the treatment of a disease or condition, refers to an amount of a compound that is capable of having any detectable, positive effect on any symptom, aspect, or characteristics of the disease or condition.

[0070] The terms "killing" or "inducing cytotoxicity" as used herein refer to inducing cell death by either apoptosis and/or necrosis, whereby embodiments of the invention include only apoptosis, only necrosis and both apoptosis and necrosis.
15

[0071] The terms "preventing" and "suppressing" as used herein refer to administering a compound prior to the onset of clinical symptoms of a disease or condition so as to prevent a physical manifestation of the disease or condition. The term "prophylaxis" is distinct from "treatment" and
20 encompasses "preventing" and "suppressing". Herein, "protection" includes "prophylaxis". Protection need not be absolute to be useful.

[0072] The term "treating" as used herein refers to administering a compound after the onset of clinical symptoms. The term "in need of treatment" as used herein refers to a judgment made by a caregiver that an individual or animal requires or will benefit from treatment. This judgment is made
25 based on a variety of factors that are in the realm of a caregiver's expertise, but that include the knowledge that the individual or animal is ill, or will be ill, as the result of a condition that is treatable by a compound of the invention.

[0073] The term "individual" or "patient" as used herein refers to any animal, including mammals, preferably mice, rats, other rodents, rabbits, dogs, cats, swine, cattle, sheep, horses, or
30 primates, and most preferably humans. The term may specify male or female or both, or exclude male or female.

[0074] As used herein, the term “non-human animal” refers to any non-human animal, including insects, birds, rodents and more usually mammals. Preferred non-human animals include: primates; farm animals such as swine, goats, sheep, donkeys, cattle, horses, chickens, rabbits; and rodents, preferably rats or mice. As used herein, the term “animal” is used to refer to any species in 5 the animal kingdom, preferably vertebrates, including birds and fish, and more preferably a mammal. Both the terms “animal” and “mammal” expressly embrace human subjects unless preceded with the term “non-human”.

[0075] As used herein, the terms “physiologically acceptable,” “pharmaceutically acceptable,” and “pharmaceutical” are interchangeable.

10

Identity Between Nucleic Acids Or Polypeptides

[0076] The terms “percentage of sequence identity” and “percentage homology” are used interchangeably herein to refer to comparisons among polynucleotides and polypeptides, and are determined by comparing two optimally aligned sequences over a comparison window, wherein the 15 portion of the polynucleotide or polypeptide sequence in the comparison window may comprise additions or deletions (i.e., gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched 20 positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity. Identity or similarity is evaluated using any of the variety of sequence comparison algorithms and programs known in the art. Such algorithms and programs include, but are by no means limited to, TBLASTN, BLASTP, FASTA, TFASTA, CLUSTALW, FASTDB (Pearson and Lipman, (1988), PNAS 85(8):2444-2448; Altschul *et al.*, 25 (1990), J. Mol. Biol. 215(3):403-410; Thompson *et al.* (1994), Nucleic Acids Res. 22(2):4673-4680; Higgins *et al.*, (1996), Meth. Enzymol. 266:383-402; Altschul *et al.*, (1993), Nature Genetics 3:266-272; Brutlag *et al.* (1990) Comp. App. Biosci. 6:237-24).

[0077] In a particularly preferred embodiment, protein and nucleic acid sequence identities are evaluated using the Basic Local Alignment Search Tool (“BLAST”) which is well known in the art 30 (e.g., Karlin and Altschul, (1990), PNAS 87:2267-2268; Altschul *et al.*, (1997), Nuc. Acids Res. 25:3389-3402). In particular, five specific BLAST programs are used to perform the following task:

- (1) BLASTP and BLAST3 compare an amino acid query sequence against a protein sequence database;
- (2) BLASTN compares a nucleotide query sequence against a nucleotide sequence database;
- 5 (3) BLASTX compares the six-frame conceptual translation products of a query nucleotide sequence (both strands) against a protein sequence database;
- (4) BLASTN compares a query protein sequence against a nucleotide sequence database translated in all six reading frames (both strands); and
- 10 (5) BLASTX compares the six-frame translations of a nucleotide query sequence against the six-frame translations of a nucleotide sequence database.

[0078] The BLAST programs identify homologous sequences by identifying similar segments, which are referred to herein as “high-scoring segment pairs,” between a query amino or nucleic acid sequence and a test sequence which is preferably obtained from a protein or nucleic acid sequence database. High-scoring segment pairs are preferably identified (i.e., aligned) by means of a scoring matrix, many of which are known in the art. Preferably, the scoring matrix used is the BLOSUM62 matrix (Gonnet *et al.*, (1992), Science 256:1443-1445; Henikoff and Henikoff, (1993), Proteins 17:49-61). Less preferably, the PAM or PAM250 matrices may also be used (see, e.g., Schwartz and Dayhoff, (1978), eds., Matrices for Detecting Distance Relationships: Atlas of Protein Sequence and Structure, Washington: National Biomedical Research Foundation). The BLAST programs evaluate the statistical significance of all high-scoring segment pairs identified, and preferably selects those segments which satisfy a user-specified threshold of significance, such as a user-specified percent homology. Preferably, the statistical significance of a high-scoring segment pair is evaluated using the statistical significance formula of Karlin (see, e.g., Karlin and Altschul, 1990). The BLAST programs may be used with the default parameters or with modified parameters 25 provided by the user.

[0079] Another preferred method for determining the best overall match between a query nucleotide or amino acid sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag *et al.* (1990). In a sequence alignment the query and subject 30 sequences are both DNA or amino acid sequences. An RNA sequence can also be compared by first converting U's to T's. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB alignment of DNA sequences to calculate percent identity are: Matrix

= Unitary, k-tuple = 4, Mismatch Penalty = 1, Joining Penalty = 30, Randomization Group Length = 0, Cutoff Score = 1, Gap Penalty = 5, Gap Size Penalty = 0.05, Window Size = 500 or the length of the subject nucleotide sequence, whichever is shorter. Preferred parameters used in a FASTDB amino acid alignment are: Matrix = PAM 0, k-tuple = 2, Mismatch Penalty = 1, Joining Penalty = 20, 5 Randomization Group Length = 0, Cutoff Score = 1, Window Size = sequence length, Gap Penalty = 5, Gap Size Penalty = 0.05, Window Size = 500 or the length of the subject amino acid sequence, whichever is shorter. If the subject sequence is shorter than the query sequence because of 5' or 3' deletions (for nucleotide sequences) or N- or C-terminal deletions (for amino acid sequences), not because of internal deletions, a manual correction must be made to the results, because the FASTDB 10 program does not account for terminal truncations of a subject sequence when calculating percent identity. The manual correction involves determining the percent of the total query sequence that is not aligned because of a truncation in the subject sequence, and subtracting this percentage from the percent identity calculated by the FASTDB program. This corrected score is what is used for the purposes of the present invention. No other manual corrections are made for the purposes of the 15 present invention.

Polynucleotides of the invention

[0080] Many of the methods described in the present specification rely on the use of common molecular biological techniques. A number of such techniques are taught, generally, in "Molecular 20 Cloning; A Laboratory Manual", 2d ed., Cole Spring Harbor Laboratory Press, Sambrook, et al., eds., 1989, and "Methods in Enzymology; Guide to Molecular Cloning Techniques", Academic Press, Berger and Kimmel eds., 1987.

[0081] The present invention concerns GENSET genomic and cDNA sequences. The present invention encompasses GENSET genes, polynucleotides comprising GENSET genomic and cDNA 25 sequences, as well as fragments and variants thereof. These polynucleotides may be purified, isolated, or recombinant.

[0082] Also encompassed by the present invention are allelic variants, orthologs, splice variants, and/or species homologues of the GENSET genes. Procedures known in the art can be used to obtain full-length genes and cDNAs, allelic variants, splice variants, full-length coding portions, 30 orthologs, and/or species homologues of genes and cDNAs corresponding to a nucleotide sequence selected from the group consisting of sequences of odd SEQ ID NOs:1-23 and sequences of clone inserts of the deposited clone pool, using information from the sequences disclosed herein or the clone

pool deposited with the ATCC. For example, allelic variants, orthologs and/or species homologues may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source for allelic variants and/or the desired homologue using any technique known to those skilled in the art including those described into the section entitled 5 "To find similar sequences".

[0083] In a specific embodiment, the polynucleotides of the invention are at least 15, 30, 50, 100, 125, 500, or 1000 continuous nucleotides. In another embodiment, the polynucleotides are less than or equal to 300kb, 200kb, 100kb, 50kb, 10kb, 7.5kb, 5kb, 2.5kb, 2kb, 1.5kb, or 1kb in length. In a further embodiment, polynucleotides of the invention comprise a portion of the coding sequences, as 10 disclosed herein, but do not comprise all or a portion of any intron. In another embodiment, the polynucleotides comprising coding sequences do not contain coding sequences of a genomic flanking gene (i.e., 5' or 3' to the gene of interest in the genome). In other embodiments, the polynucleotides of the invention do not contain the coding sequence of more than 1000, 500, 250, 100, 75, 50, 25, 20, 15, 10, 5, 4, 3, 2, or 1 naturally occurring genomic flanking gene(s).

15 Deposited clone pool of the invention

[0084] Expression of GENSET genes has been shown to lead to the production of at least one mRNA species per GENSET gene, which cDNA sequence is set forth in the appended Sequence Listing as odd SEQ ID NOS:1-23. The cDNAs corresponding to these GENSET mRNA species were cloned either in the vector pBluescriptII SK⁻ (Stratagene) or in a vector called pPT. Cells containing 20 the cloned cDNAs of the present invention are maintained in permanent deposit by the inventors at Genset, S.A., 24 Rue Royale, 75008 Paris, France. Table I provides Genset's internal designation number assigned to each SEQ ID NO., and indicates whether the sequence is a nucleic acid sequence (DNA) or a protein (PRT) sequence. Each cDNA can be removed from the Bluescript vector in which it was inserted by performing a NotI Pst I double digestion, or from the pPT vector by performing a 25 MunI HindIII double digestion, to produce the appropriate fragment for each clone, provided the cDNA sequence does not contain any of the corresponding restriction sites within its sequence. Alternatively, other restriction enzymes of the multicloning site of the vector may be used to recover the desired insert as indicated by the manufacturer.

[0085] Pools of cells containing GENSET genes as described in the Sequence Listing, from 30 which the cells containing a particular polynucleotide is obtainable, were or will be also deposited with the American Tissue Culture Collection (ATCC), 10801 University Boulevard, Manassas, VA 20110-

2209, United States. Each cDNA clone has been transfected into separate bacterial cells (E-coli) for these composite deposits.

[0086] Bacterial cells containing a particular clone can be obtained from the composit deposit using standard methods, e.g., by plating a culture of the composit deposit, transferring the lysed 5 colonies to a filter, and hybridizing the filter with a labelled probe such as an oligonucleotide probe specific for the clone of interest. Alternatively, individual cDNA inserts can be recovered from the pool of bacteria using PCR with primers designed at both ends of the cDNA insertion, including primers designed in the multicloning site of the vector.

cDNA sequences of the invention

10 [0087] Structural parameters of each of the cDNAs of the present invention are presented in the appended Sequence Listing. Accordingly, the coding sequence (CDS) or open reading frame (ORF) of each cDNA of the invention refers to the nucleotide sequence beginning with the first nucleotide of the start codon and ending with the nucleotide immediately 5' to the first nucleotide of the stop codon. Similarly, the 5' untranslated region (or 5'UTR) of each cDNA of the invention refers 15 to the nucleotide sequence starting at nucleotide 1 and ending at the nucleotide immediately 5' to the first nucleotide of the start codon. The 3' untranslated region (or 3'UTR) of each cDNA of the invention refers to the nucleotide sequence starting at the first nucleotide of the stop codon and ending at the last nucleotide of the cDNA.

Coding sequences

20 [0088] Another object of the invention is an isolated, purified or recombinant polynucleotide comprising the coding sequence of a sequence selected from the group consisting of the polynucleotide sequences of the appended Sequence Listing, those of human cDNA clone inserts of the deposited clone pool and variants thereof.

[0089] It will be appreciated that should the extent of the coding sequence differ from that 25 indicated in the appended sequence listing as a result of a sequencing error, reverse transcription or amplification error, mRNA splicing, post-translational modification of the encoded protein, enzymatic cleavage of the encoded protein, or other biological factors, one skilled in the art would be readily able to identify the extent of the coding sequences in the polynucleotide sequences of the Sequence Listing, those of the human cDNA inserts of the deposited clone pool, and allelic variants thereof. 30 Accordingly, the scope of any claims herein relating to nucleic acids containing the coding sequence of one of the polynucleotide sequences of the Sequence Listing and those of the cDNA inserts of the deposited clone pool is not to be construed as excluding any readily identifiable variations from or

equivalents to the coding sequences described in the appended sequence listing. Equivalents include any alterations in a nucleotide coding sequence that does not result in an amino acid change, or that results in a conservative amino acid substitution, as defined below, in the polypeptide encoded by the nucleotide sequence. Similarly, should the extent of the polypeptides differ from those indicated in the 5 appended Sequence Listing as a result of any of the preceding factors, the scope of claims relating to polypeptides comprising the amino acid sequence of the polypeptide sequences of the appended Sequence Listing is not to be construed as excluding any readily identifiable variations from or equivalents to the sequences described in the appended sequence listing.

[0090] The above disclosed polynucleotides that contain the coding sequence of the GENSET genes may be expressed in a desired host cell or a desired host organism, when this polynucleotide is placed under the control of suitable expression signals. The expression signals may be either the expression signals contained in the regulatory regions in the GENSET genes of the invention or, in contrast, the signals may be exogenous regulatory nucleic sequences. Such a polynucleotide, when placed under the suitable expression signals, may also be inserted in a vector for its expression and/or 15 amplification.

[0091] Further included in the present invention are polynucleotides encoding the polypeptides of the present invention that are fused in frame to the coding sequences for additional heterologous amino acid sequences. Also included in the present invention are nucleic acids encoding polypeptides of the present invention together with additional, non-coding sequences, including, but 20 not limited to, non-coding 5' and 3' sequences, vector sequence, sequences used for purification, probing, or priming. For example, heterologous sequences include transcribed, untranslated sequences that may play a role in transcription and mRNA processing, such as ribosome binding and stability of mRNA. The heterologous sequences may alternatively comprise additional coding sequences that provide additional functionalities, e.g. a hexahistidine or HA tag).

25 Regulatory sequences of the invention

[0092] As mentioned, the genomic sequence of GENSET genes contain regulatory sequences in the non-coding 5'-flanking region and possibly in the non-coding 3'-flanking region that border the GENSET polypeptide coding regions containing the exons of these genes.

[0093] Polynucleotides derived from GENSET polynucleotide 5' and 3' regulatory regions 30 are useful in order to detect the presence of at least a copy of a genomic nucleotide sequence of the GENSET gene or a fragment thereof in a test sample.

Preferred regulatory sequences

[0094] Polynucleotides carrying the regulatory elements located at the 5' end and at the 3' end of GENSET polypeptide coding regions may be advantageously used to control the processing, localization, stability, maturation and transcriptional and translational activity of a heterologous 5 polynucleotide of interest, e.g; the regulatory polynucleotides may be part of a recombinant expression vector that may be used to express a coding sequence in a desired host cell or host organism (for a review on UTRs see Decker and Parker, (1995) Curr. Opin. Cell. Biol. 7(3):368-92, Derrigo *et al.*, (2000) Int. J. Mol. Med. 5(2):111-23). In particular, 3'UTRs may be used in order to control the stability of heterologous mRNAs in recombinant vectors using any methods known to those skilled in 10 the art including Makrides (1999) Protein Expr Purif 1999 Nov; 17(2):183-202), US Patents 5,925,564; 5,807,707 and 5,756,264.

[0095] The present invention also concerns a purified or isolated nucleic acid comprising a polynucleotide which is selected from the group consisting of the 5' and 3' GENSET polynucleotide regulatory regions, sequences complementary thereto, regulatory active fragments and variants thereof, 15 and those that hybridize under stringent hybridization conditions therewith. Further included are nucleic acids comprising a nucleotide sequence having at least 95% identity with any of the herein-described GENSET 5' or 3' regulatory sequences, as well as 5'- or 3'-UTRs of the polynucleotide sequences of the appended Sequence Listing, of human cDNA clone inserts of the deposited clone pool, sequences complementary thereto, regulatory active fragments and allelic variants thereof. 20 Fragments of 5' and 3' regulatory regions may have a length corresponding to any one integer between 20 and 20,000 nucleotides in length.

[0096] For the purpose of the invention, a nucleic acid or polynucleotide is "functional" as a "regulatory region" for expressing a recombinant polypeptide or a recombinant polynucleotide if said regulatory polynucleotide contains nucleotide sequences which contain transcriptional and 25 translational regulatory information, and such sequences are "operably linked" to nucleotide sequences which encode the desired polypeptide or the desired polynucleotide.

[0097] The invention also comprises a nucleic acid molecule encoding a desired polypeptide or a nucleic acid molecule of interest, wherein said nucleic acid molecule is operably linked to any of the herein-described regulatory sequences. The desired polypeptide may be of various nature or 30 origin, encompassing proteins of prokaryotic viral or eukaryotic origin. Also encompassed are eukaryotic proteins such as intracellular proteins, such as "house keeping" proteins, membrane-bound proteins, such as mitochondrial membrane-bound proteins and cell surface receptors, and secreted

proteins such as endogenous mediators such as cytokines. The desired polypeptide may be a heterologous polypeptide or a GENSET polypeptide.

Polynucleotide variants

[0098] The invention also relates to variants of the polynucleotides described herein and 5 fragments thereof. "Variants" of polynucleotides, as the term is used herein, are polynucleotides that differ from a reference polynucleotide. Generally, differences are limited so that the nucleotide sequences of the reference and the variant are closely similar overall and, in many regions, identical. The present invention encompasses both allelic variants and degenerate variants.

Allelic variant

10 [0099] A variant of a polynucleotide may be a naturally occurring variant such as a naturally occurring allelic variant, or it may be a variant that is not known to occur naturally. By an "allelic variant" is intended one of several alternate forms of a gene occupying a given locus on a chromosome of an organism (see Lewin, (1989), PNAS 86:9832-8935). Diploid organisms may be homozygous or heterozygous for an allelic form. Non-naturally occurring variants of the polynucleotide may be made 15 by art-known mutagenesis techniques, including those applied to polynucleotides, cells or organisms.

Degenerate variant

[0100] In addition to the isolated polynucleotides of the present invention, and fragments thereof, the invention further includes polynucleotides which comprise a sequence substantially different from those described above but which, due to the degeneracy of the genetic code, still encode 20 a GENSET polypeptide of the present invention. These polynucleotide variants are referred to as "degenerate variants" throughout the instant application. That is, all possible polynucleotide sequences that encode the GENSET polypeptides of the present invention are contemplated. This includes the genetic code and species-specific codon preferences known in the art.

[0101] Nucleotide changes present in a variant polynucleotide may be silent, which means 25 that they do not alter the amino acids encoded by the polynucleotide. However, nucleotide changes may also result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence. The substitutions, deletions or additions may involve one or more nucleotides. The variants may be altered in coding or non-coding regions or both. Alterations in the coding regions may produce conservative or non-conservative amino acid 30 substitutions, deletions or additions. In the context of the present invention, preferred embodiments are those in which the polynucleotide variants encode polypeptides which retain substantially the same

biological properties or activities as the GENSET protein. More preferred polynucleotide variants are those containing conservative substitutions.

Similar polynucleotides

[0102] Other embodiments of the present invention provide a purified, isolated or recombinant polynucleotide which is at least 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to a polynucleotide of the present invention. Although similar polynucleotides encoding polypeptides with GENSET biological activity are preferred, the presence of GENSET biological activity in an encoded protein is not necessary because even where a particular nucleic acid molecule does not encode a polypeptide having activity, one of skill in the art would still know how to use the nucleic acid molecule, for instance, as a hybridization probe or primer. Uses of the nucleic acid molecules of the present invention that do not encode a polypeptide having GENSET activity include, inter alia, isolating a GENSET gene or allelic variants thereof from a DNA library, and detecting GENSET mRNA expression in biological samples suspected of containing GENSET mRNA or DNA, e.g., by Northern Blot or PCR analysis.

15 Hybridizing Polynucleotides

[0103] In another aspect, the invention provides an isolated or purified nucleic acid molecule comprising a polynucleotide which hybridizes under stringent hybridization conditions to any polynucleotide of the present. Such hybridizing polynucleotides may be of at least any one integer between 10 and 10,000 nucleotides in length.

20 [0104] Of course, a polynucleotide which hybridizes only to polyA+ sequences (such as any 3' terminal polyA+ tract of a cDNA shown in the sequence listing), or to a 5' complementary stretch of T (or U) residues, would not be included in the definition of "polynucleotide," since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly(A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone generated using oligo dT as a 25 primer).

Complementary polynucleotides

[0105] The invention further provides isolated nucleic acid molecules having a nucleotide sequence fully complementary to any polynucleotide of the invention.

Polynucleotide fragments

30 [0106] The present invention is further directed to portions or fragments of the polynucleotides of the present invention. Uses for the polynucleotide fragments of the present

invention include probes, primers, molecular weight markers and for expressing the polypeptide fragments of the present invention. Fragments include portions of polynucleotides selected from the group consisting of a) polynucleotide sequences of the Sequence Listing, b) genomic GENSET sequences, c) polynucleotides encoding a polypeptide of the present invention, d) sequences of human 5 cDNA clone inserts of the deposited clone pool, and e) polynucleotides encoding the polypeptides encoded by the human cDNA clone inserts of the deposited clone pool. Particularly included in the present invention is a purified or isolated polynucleotide comprising at least 8, 10, 12, 15, 18, 20, 25, 28, 30, 35, 40, 50, 75, 100, 150, 200, 300, 400, 500, 800, 1000, 1500, or 2000 consecutive nucleotides of a polynucleotide of the present invention. Further, polynucleotides are provided comprising at least 10 X nucleotides, wherein "X" is defined as any integer between 8 and the integer representing the 3' most nucleotide position as set forth in the sequence listing or elsewhere herein.

[0107] Further included as preferred polynucleotides of the present invention are polynucleotide fragments that are specified in terms of their 5' and 3' position. Where the 5' and 3' positions are represented by the position numbers set forth in the appended sequence listing wherein 15 the 5' most nucleotide is 1 and the 3' most nucleotide is the last nucleotide for a particular SEQ ID No., all polynucleotide fragments corresponding to every combination of a 5' and 3' nucleotide position that a polynucleotide fragment of the present invention, at least 8 contiguous nucleotides in length, could occupy on a polynucleotide of the invention are specifically considered. These species of polynucleotide fragments may alternatively be described by the formula "a to b"; where "a" equals the 20 5' most nucleotide position and "b" equals the 3' most nucleotide position of the polynucleotide; and further where "a" equals an integer between 1 and the number of nucleotides of the polynucleotide sequence of the present invention minus 8, and where "b" equals an integer between 9 and the number of nucleotides of the polynucleotide sequence of the present invention; and where "a" is an integer smaller than "b" by at least 8. All of the polynucleotide fragments described in either of these ways 25 can be immediately envisaged and are therefore not individually listed solely for the purpose of not unnecessarily lengthening the specification. Any of these polynucleotide fragments may also be specifically excluded from the present invention. Any number of fragments specified by 5' and 3' positions or by size in nucleotides, as described above, may be excluded. Preferred excluded fragments include those having substantial homology to repeated sequences including Alu, L1, THE 30 and MER repeats, SSTR sequences or satellite, micro-satellite, and telomeric repeats.

[0108] Other preferred fragments of the invention are polynucleotides comprising polynucleotide sequences encoding domains of polypeptides. Such fragments may be used to obtain other polynucleotides encoding polypeptides having similar domains using hybridization or RT-PCR

techniques. Alternatively, these fragments may be used to express a polypeptide domain which may have a specific biological property.

[0109] Another object of the invention is an isolated, purified or recombinant polynucleotide encoding a polypeptide consisting of, consisting essentially of, or comprising a contiguous span of at least (any integer between 5 and 1,000 consecutive amino acids in length more preferably at least) 5, 6, 8, 10, 12, 15, 20, 25, 30, 35, 40, 50, 60, 75, 100, 150 or 200 consecutive amino acids.

[0110] The present invention further encompasses any combination of the polynucleotide fragments listed in this section.

Oligonucleotide primers and probes

10 [0111] The present invention also encompasses fragments of GENSET polynucleotides for use as primers and probes. Polynucleotides derived from the GENSET genomic and cDNA sequences are useful in order to detect the presence of at least a copy of a GENSET polynucleotide or fragment, complement, or variant thereof in a test sample.

Structural definition

15 [0112] Any polynucleotide of the invention may be used as a primer or probe. Particularly preferred probes and primers of the invention include isolated, purified, or recombinant polynucleotides comprising a contiguous span of at least 12, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500, or 1000 nucleotides of a polynucleotide of the present invention.

20 [0113] For amplification purposes, pairs of primers with approximately the same Tm are preferable. Primers may be designed using methods known in the art. Amplification techniques that can be used in the context of the present invention include, but are not limited to, the ligase chain reaction (LCR) described in EP-A- 320 308, WO 9320227 and EP-A-439 182, the polymerase chain reaction (PCR, RT-PCR) and techniques such as the nucleic acid sequence based amplification (NASBA) described in Guatelli *et al.*, (1990) PNAS 35:273-286 and in Compton (1991) Nature 25 350(6313):91-92, Q-beta amplification as described in European Patent Application No 4544610, strand displacement amplification as described in Walker, *et al.* (1996), Clin. Chem. 42:9-13 and EP A 684 315 and, target mediated amplification as described in WO 9322461.

30 [0114] The probes of the present invention are useful for a number of purposes, including for Southern hybridization to genomic DNA, to detect PCR amplification products, to detect mismatches in a GENSET gene or mRNA, and in in situ hybridization. Any of the polynucleotides, primers and probes of the present invention can be conveniently immobilized on any type of solid support, such as

latex particles, microparticles, magnetic beads, non-magnetic beads (including polystyrene beads), membranes (including nitrocellulose strips), plastic tubes, walls of microtiter wells, glass or silicon chips, sheep (or other suitable animal's) red blood cells and duracytes. Suitable methods for immobilizing nucleic acids on solid phases include ionic, hydrophobic, covalent interactions and the like. A solid support, as used herein, refers to any material which is insoluble, or can be made insoluble by a subsequent reaction. The solid support can be chosen for its intrinsic ability to attract and immobilize the capture reagent, or alternatively may retain an additional receptor which has the ability to attract and immobilize the capture reagent. The polynucleotides of the invention can be attached to or immobilized on a solid support individually or in groups of at least 2, 5, 8, 10, 12, 15, 20, or 25 distinct polynucleotides of the invention to a single solid support. In addition, polynucleotides other than those of the invention may be attached to the same solid support as one or more polynucleotides of the invention.

Oligonucleotide array

[0115] A substrate comprising a plurality of oligonucleotide primers or probes of the invention may be used either for detecting or amplifying targeted sequences in GENSET genes, may be used for detecting mutations in the coding or in the non-coding sequences of GENSET genes, and may also be used to determine GENSET gene expression in different contexts such as in different tissues, at different stages of a process (embryo development, disease treatment), and in patients versus healthy individuals as described elsewhere in the application.

[0116] As used herein, the term "array" means a one dimensional, two dimensional, or multidimensional arrangement of nucleic acids of sufficient length to permit specific detection of gene expression. For example, the array may contain a plurality of nucleic acids derived from genes whose expression levels are to be assessed. The array may include any of the herein-described GENSET genomic DNA, a GENSET cDNA, sequences complementary thereto or fragments thereof. Preferably, the fragments are at least 12, 15, 18, 20, 25, 30, 35, 40, 50, or 100 nucleotides in length.

[0117] Any polynucleotide provided herein may be attached in overlapping areas or at random locations on the solid support. Alternatively the polynucleotides of the invention may be attached in an ordered array wherein each polynucleotide is attached to a distinct region of the solid support which does not overlap with the attachment site of any other polynucleotide. Preferably, such an ordered array of polynucleotides is designed to be "addressable" where the distinct locations are recorded and can be accessed as part of an assay procedure. Addressable polynucleotide arrays typically comprise a plurality of different oligonucleotide probes that are coupled to a surface of a

substrate in different known locations. The knowledge of the precise location of each polynucleotides location makes these “addressable” arrays particularly useful in hybridization assays. Any addressable array technology known in the art can be employed with the polynucleotides of the invention, including Genechips™ (see, e.g., US Patent No. 5,143,854; PCT publications WO 90/15070 and 5 92/10092, Fodor *et al.*, (1991) *Science* 251:767-777). The immobilization of arrays of oligonucleotides on solid supports has been rendered possible by the development of a technology generally identified as “Very Large Scale Immobilized Polymer Synthesis” (VLSIPS™) in which, typically, probes are immobilized in a high density array on a solid surface of a chip (see, e.g., US Patents 5,143,854 and 5,412,087, and in PCT Publications WO 90/15070, WO 92/10092 and WO 10 95/11995). Further presentation strategies known in the art may be used, such as those disclosed in WO 94/12305, WO 94/11530, WO 97/29212 and WO 97/31256.

[0118] Consequently, the invention concerns an array of nucleic acid molecules comprising at least one, two, or five polynucleotides of the invention, particularly probes or primers as described herein.

15

Methods of making the polynucleotides of the invention

[0119] The present invention also comprises methods of making the polynucleotides of the invention. Polynucleotides of the invention may be synthesized either enzymatically using techniques well known to those skilled in the art including amplification or hybridization-based methods as described herein, or chemically.

20

[0120] A variety of chemical methods of synthesizing nucleic acids are known to those skilled in the art. In many of these methods, synthesis is conducted on a solid support. Alternatively, polynucleotides may be prepared as described in U.S. Patent No. 5,049. In some embodiments, several polynucleotides prepared as described above are ligated together to generate longer polynucleotides having a desired sequence.

25

Polypeptides of the invention

[0121] The term “GENSET polypeptides” is used herein to embrace all of the proteins and polypeptides of the present invention. The present invention encompasses GENSET polypeptides, including recombinant, isolated or purified GENSET polypeptides consisting of: (a) the full length polypeptides of even SEQ ID NOs:2-24; (b) the full length polypeptides encoded by the clone inserts 30 of the deposited clone pool; (c) the epitope-bearing fragments of the polypeptides of even SEQ ID NOs:2-24; (d) the epitope-bearing fragments of the polypeptides encoded by the clone inserts contained in the deposited clone pool; (e) the domains of the polypeptides of even SEQ ID NOs:2-24;

(f) the domains of the polypeptides encoded by the clone inserts contained in the deposited clone pool;
(g) the signal peptides of the polypeptides of even SEQ ID NOS:2-24 or encoded by the human cDNAs of the deposited clone pool; (h) the mature polypeptides of even SEQ ID NOS:2-24 or encoded by the human cDNAs of the deposited clone pool; and (i) the allelic variant polypeptides of any of the 5 polypeptides of (a)-(f). Other objects of the invention are polypeptides encoded by the polynucleotides of the invention as well as fusion polypeptides comprising such polypeptides.

Polypeptide variants

[0122] The present invention further provides for GENSET polypeptides encoded by allelic and splice variants, orthologs, and/or species homologues. Procedures known in the art can be used to 10 obtain, allelic variants, splice variants, orthologs, and/or species homologues of polynucleotides encoding polypeptides of the Sequence Listing and polypeptides encoded by the clone inserts of the deposited clone pool, using information from the sequences disclosed herein or the clones deposited with the ATCC.

[0123] The polypeptides of the present invention also include polypeptides having an amino acid sequence at least 50% identical, more preferably at least 60% identical, and still more preferably 15 70%, 80%, 90%, 95%, 96%, 97%, 98% or 99% identical to a polypeptide of the present invention. By a polypeptide having an amino acid sequence at least, for example, 95% "identical" to a query amino acid sequence of the present invention, it is intended that the amino acid sequence of the subject polypeptide is identical to the query sequence except that the subject polypeptide sequence may 20 include up to five amino acid alterations per each 100 amino acids of the query amino acid sequence. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a query amino acid sequence, up to 5% (5 of 100) of the amino acid residues in the subject sequence may be inserted, deleted, (indels) or substituted with another amino acid.

[0124] Further polypeptides of the present invention include polypeptides which have at least 25 90% similarity, more preferably at least 95% similarity, and still more preferably at least 96%, 97%, 98% or 99% similarity to those described above. "Similarity" is calculated exactly as described above for "identity", except that for the purposes of the calculation a matching amino acid can be either identical or an amino acid representing an "equivalent" change, as defined below.

[0125] These alterations of the reference sequence may occur at the amino or carboxy 30 terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence. The variant polypeptides described herein are included in the

present invention regardless of whether they have their normal biological activity, as one of skill in the art would recognize that variant polypeptides lacking biological activity would still be useful, for instance, as a vaccine, to generate antibodies, as epitope tags, in epitope mapping, as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns.

5 Preparation of the polypeptides of the invention

[0126] The polypeptides of the present invention can be prepared in any suitable manner known in the art. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. The polypeptides of the present invention are preferably provided in an 10 isolated form, and may be partially or preferably substantially purified.

Isolation

From natural sources

[0127] The GENSET proteins of the invention may be isolated from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured cells, of humans or non-15 human animals. Methods for extracting and purifying natural proteins are known in the art, and include the use of detergents or chaotropic agents to disrupt particles followed by differential extraction and separation of the polypeptides by ion exchange chromatography, affinity chromatography, sedimentation according to density, and gel electrophoresis. See, e.g., "Methods in Enzymology, Academic Press, 1993" for a variety of methods for purifying proteins. Polypeptides of 20 the invention also can be purified from natural sources using antibodies directed against the polypeptides of the invention, using standard methods.

From recombinant sources

[0128] Preferably, the GENSET polypeptides of the invention are recombinantly produced using routine expression methods known in the art. The polynucleotide encoding the desired 25 polypeptide is operably linked to a promoter into an expression vector suitable for any convenient host. Both eukaryotic and prokaryotic host systems are used in forming recombinant polypeptides. The polypeptide is then isolated from lysed cells or from the culture medium and purified to the extent needed for its intended use. Any polynucleotide of the present invention may be used to express GENSET polypeptides.

30 [0129] Consequently, a further embodiment of the present invention is a method of making a polypeptide of the present invention, said method comprising the steps of providing a GENSET

polynucleotide (e.g. a polynucleotide encoding a GENSET polypeptide), inserting the polynucleotide in an expression vector such that the cDNA is operably linked to a promoter; and introducing the expression vector into a host cell whereby said host cell produces said polypeptide.

[0130] In one aspect of this embodiment, the method further comprises the step of isolating 5 the polypeptide. Any suitable expression vector and system (e.g. cell-based system such as 3T3 cells) may be used, according to methods well known in the art.

[0131] In one embodiment, the entire coding sequence of a GENSET cDNA and the 3'UTR through the polyA signal of the cDNA is operably linked to a promoter in the expression vector.

[0132] In another embodiment, an additional nucleotide sequence is included which codes for 10 secretory or leader sequences, pro-sequences, sequences which aid in purification, such as multiple histidine residues, or an additional sequence for stability during recombinant production.

[0133] Alternatively, the GENSET polypeptide to be expressed may also be a product of transgenic animals, i.e., as a component of the milk of transgenic cows, goats, pigs or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein of 15 interest.

[0134] Any standard method may be used to recover a GENSET polypeptide expressed using these methods, including differential extraction, ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, antibody-based methods, affinity chromatography, hydroxylapatite 20 chromatography, HPLC, immunochromatography, and lectin chromatography.

[0135] Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated, and may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. In addition, polypeptides of the present invention may or may not contain the amino terminal 25 methionine.

From chemical synthesis

[0136] In addition, polypeptides of the invention, especially short protein fragments, can be chemically synthesized using techniques known in the art (*See, e.g., Creighton (1983), Proteins: Structures and Molecular Principles, W.H. Freeman & Co. 2nd Ed., T. E., New York; and Hunkapiller 30 et al., (1984) Nature. 310(5973):105-11).* For example, a polypeptide corresponding to a fragment of a

polypeptide sequence of the invention can be synthesized by use of a peptide synthesizer. Alternatively, the methods described in U.S. Patent No. 5,049,656, may be used.

[0137] Furthermore, if desired, nonclassical amino acids or chemical amino acid analogs can be introduced as a substitution or addition into the polypeptide sequence. Non-classical amino acids include, but are not limited to, to the D-isomers of the common amino acids, 2,4-diaminobutyric acid, α-amino isobutyric acid, 4-aminobutyric acid, Abu, 2-amino butyric acid, g-Abu, e-Ahx, 6-amino hexanoic acid, Aib, 2-amino isobutyric acid, 3-amino propionic acid, ornithine, norleucine, norvaline, hydroxyproline, sarcosine, citrulline, homocitrulline, cysteic acid, t-butylglycine, t-butylalanine, phenylglycine, cyclohexylalanine, b-alanine, fluoroamino acids, designer amino acids such as b-methyl amino acids, Ca-methyl amino acids, Na-methyl amino acids, and amino acid analogs in general. Furthermore, the amino acid can be D (dextrorotary) or L (levorotary).

Modifications

[0138] The invention encompasses polypeptides which are differentially modified during or after translation, e.g., by glycosylation, acetylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to an antibody molecule or other cellular ligand, etc. Any of numerous chemical modifications may be carried out by known techniques, including, but not limited to, specific chemical cleavage by cyanogen bromide, trypsin, chymotrypsin, papain, V8 protease, NaBH4; acetylation, formylation, oxidation, reduction; metabolic synthesis in the presence of tunicamycin; etc.

[0139] Additional post-translational modifications encompassed by the invention include, for example, e.g., N-linked or O-linked carbohydrate chains, processing of N-terminal or C-terminal ends, attachment of chemical moieties to the amino acid backbone, chemical modifications of N-linked or O-linked carbohydrate chains, and addition or deletion of an N-terminal methionine residue as a result of prokaryotic host cell expression. The polypeptides may also be modified with a detectable label, such as an enzymatic, fluorescent, isotopic or affinity label to allow for detection and isolation of the protein.

[0140] Also provided by the invention are chemically modified derivatives of the polypeptides of the invention which may provide additional advantages such as increased solubility, stability and circulating time of the polypeptide, or decreased immunogenicity (see, e.g., U.S. Patent No. 4,179,337). The chemical moieties for derivatization may be selected from water soluble polymers (branched or unbranched) such as polyethylene glycol (preferably between 1 and 100 kD), ethylene glycol/propylene glycol copolymers, carboxymethylcellulose, dextran, polyvinyl alcohol and

the like. The polypeptides may be modified at random positions within the molecule, or at predetermined positions within the molecule and may include one, two, three or more attached chemical moieties. The chemical moiety may be attached using any method, e.g. via a free amino, carboxyl, or sulphydryl group (see, e.g., EP 0 401 384, or Malik *et al.*, (1992), Exp. Hematol. 5 20:1028-1035).

Multimerization

[0141] The polypeptides of the invention may be in monomers or multimers (i.e., dimers, trimers, tetramers and higher multimers). Accordingly, the present invention relates to monomers and multimers of the polypeptides of the invention, their preparation, and compositions containing them.

10 [0142] Multimers encompassed by the invention may be homomers or heteromers. As used herein, the term "homomer" refers to a multimer containing only polypeptides with the same amino acid sequence (although a small amount of variation is allowed), and "heteromer" refers to a multimer containing one or more heterologous polypeptides.

[0143] Multimers of the invention may be the result of hydrophobic, hydrophilic, ionic and/or covalent associations and/or may be indirectly linked. Thus, in one embodiment, multimers of the invention, such as, for example, homodimers or homotrimers, are formed when polypeptides of the invention contact one another in solution. In another embodiment, heteromultimers of the invention, such as, for example, heterotrimers or heterotetramers, are formed when polypeptides of the invention contact antibodies to the polypeptides of the invention (including antibodies to the heterologous 20 polypeptide sequence in a fusion protein of the invention) in solution. In other embodiments, multimers of the invention are formed by cross-linking between cysteine residues located within the polypeptide sequences. Any of these associations may involve one or more amino acid residues contained in an amino acid provided in the sequence listing or in a heterologous polypeptide sequence of a fusion protein, e.g., in an Fc fusion protein, osteoprotegerin fusion protein, peptide linker fusion 25 protein, Flag® fusion protein, or leucine zipper fusion protein of the invention (see, e.g., US Patent Number 5,478,925, WO 98/49305, or U.S. Pat. No. 5,073,627, Landschulz *et al.*, (1988), Science. 240:1759, WO 94/10308, Hoppe *et al.*, (1994), FEBS Letters. 344:191 and in U.S. patent application Ser. No. 08/446,922). Other methods of making multimers include the addition of cysteine or biotin to the C-terminus or N-terminus of the polypeptide using techniques known in the art, or by generating 30 liposomes containing the polypeptide components desired to be contained in the multimer of the invention (see, e.g., US Patent Number 5,478,925 for numerous methods of multimerization).

[0144] Multimers of the invention may be generated using chemical or genetic engineering techniques known in the art.

Mutated polypeptides

[0145] To improve or alter the characteristics of GENSET polypeptides of the present invention, recombinant DNA technology can be used to create novel mutant proteins or muteins including single or multiple amino acid substitutions, deletions, additions, or fusion proteins. Such modified polypeptides can show, e.g., increased/decreased biological activity or increased/decreased stability. In addition, they may be purified in higher yields and show better solubility than the corresponding natural polypeptide, at least under certain purification and storage conditions.

10 N- and C-terminal deletions

[0146] It is known in the art that one or more amino acids may be deleted from the N-terminus or C-terminus without substantial loss of biological function. (See, e.g., Ron *et al.*, (1993), Biol Chem., 268 2984-2988 ; Dobeli, *et al.* 1988.) Accordingly, the present invention provides polypeptides having one or more residues deleted from the amino and/or carboxy terminus.

15 Other mutations

[0147] The invention includes numerous variations of the GENSET polypeptides which show substantial GENSET polypeptide activity. Such mutants include deletions, insertions, inversions, repeats, and substitutions selected according to general rules known in the art so as to have little effect on activity.

20 [0148] There are two main approaches for studying the tolerance of an amino acid sequence to change (see, Bowie *et al.*, (1994), *Science*. 247:1306-1310). The first method relies on the process of evolution, in which mutations are either accepted or rejected by natural selection. The second approach uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene and selections or screens to identify sequences that maintain functionality. Examples of this
25 include site-directed mutagenesis or alanine-scanning mutagenesis (see, e.g., Cunningham *et al.* (1989), *Science* 244:1081-1085). These studies have revealed that proteins are surprisingly tolerant of amino acid substitutions.

30 [0149] Typically seen as conservative substitutions are the replacements, one for another, among the aliphatic amino acids Ala, Val, Leu and Phe; interchange of the hydroxyl residues Ser and Thr, exchange of the acidic residues Asp and Glu, substitution between the amide residues Asn and Gln, exchange of the basic residues Lys and Arg and replacements among the aromatic residues Phe,

Tyr. Thus, the polypeptide of the present invention may be, for example, one in which one or more of the amino acid residues are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue). Such substituted amino acid residue may or may not be one encoded by the genetic code, and may include a substituent group.

5 [0150] As indicated, changes are preferably of a minor nature, such as conservative amino acid substitutions that do not significantly affect the folding or activity of the protein. The following groups of amino acids represent equivalent changes: (1) Ala, Pro, Gly, Glu, Asp, Gln, Asn, Ser, Thr; (2) Cys, Ser, Tyr, Thr; (3) Val, Ile, Leu, Met, Ala, Phe; (4) Lys, Arg, His; (5) Phe, Tyr, Trp, His.

[0151] Furthermore, GENSET polypeptides of the present invention may include one or more 10 amino acid substitutions that mimic modified amino acids. An example of this type of substitution includes replacing amino acids that are capable of being phosphorylated (e.g., serine, threonine, or tyrosine) with a negatively charged amino acid that resembles the negative charge of the phosphorylated amino acid (e.g., aspartic acid or glutamic acid). Also included is substitution of 15 amino acids that are capable of being modified by hydrophobic groups (e.g., arginine) with amino acids carrying bulky hydrophobic side chains, such as tryptophan or phenylalanine. Therefore, a specific embodiment of the invention includes GENSET polypeptides that include one or more amino acid substitutions that mimic modified amino acids at positions where amino acids that are capable of being modified are normally positioned. Furthermore, any GENSET polypeptide amino acid capable of being modified may be excluded from substitution with a modification-mimicking amino acid.

20 [0152] A specific embodiment of a modified GENSET peptide molecule of interest according to the present invention, includes, but is not limited to, a peptide molecule which is resistant to proteolysis, is a peptide in which the -CONH- peptide bond is modified and replaced by a (CH₂NH) reduced bond, a (NHCO) retro inverso bond, a (CH₂-O) methylene-oxy bond, a (CH₂-S) thiomethylene bond, a (CH₂CH₂) carba bond, a (CO-CH₂) cetomethylene bond, a (CHOH-CH₂) 25 hydroxyethylene bond), a (N-N) bound, a E-alcene bond or also a -CH=CH- bond.

[0153] Of special interest are substitutions of charged amino acids with other charged or neutral amino acids which may produce proteins with highly desirable improved characteristics, such as less aggregation. Aggregation may not only reduce activity but also be problematic when preparing pharmaceutical formulations, because aggregates can be immunogenic (*see, e.g.,* Pinckard *et al.*, 30 (1967), Clin. Exp. Immunol. 2:331-340; Robbins *et al.*, (1987), Diabetes. 36:838-845; and Cleland *et al.*, (1993) Crit. Rev. Ther. Drug Carr. Syst. 10:307-377).

[0154] A further embodiment of the invention relates to a polypeptide which comprises the amino acid sequence of a GENSET polypeptide having an amino acid sequence which contains at least any one integer from 1 to 50 of conservative amino acid substitutions. Any conservative substitution or combination of substitutions may also be excluded.

5 Polypeptide fragments

Structural definition

[0155] The present invention is further directed to fragments of the polypeptides of the present invention. More specifically, the present invention embodies purified, isolated, and recombinant polypeptides comprising at least any one integer between 6 and 1000 (or the length of the 10 polypeptides amino acid residues minus 1 if the length is less than 1000) of consecutive amino acid residues. Preferably, the fragments are at least 6, preferably at least 8 to 10, more preferably 12, 15, 20, 25, 30, 35, 40, 50, 60, 75, 100, 125, 150, 175, 200, 225, 250, 275, or 300 consecutive amino acids of a polypeptide of the present invention.

[0156] In addition to the above polypeptide fragments, further preferred sub-genuses of 15 polypeptides comprise at least X amino acids, wherein "X" is defined as any integer between 6 and the integer representing the C-terminal amino acid of the polypeptide of the present invention including the polypeptide sequences of the sequence listing below. Further included are species of polypeptide fragments at least 6 amino acids in length, as described above, that are further specified in terms of their N-terminal and C-terminal positions. However, included in the present invention as individual 20 species are all polypeptide fragments, at least 6 amino acids in length, as described above, and may be particularly specified by a N-terminal and C-terminal position. That is, every combination of a N-terminal and C-terminal position that a fragment at least 6 contiguous amino acid residues in length could occupy, on any given amino acid sequence of the sequence listing or of the present invention is included in the present invention

25 [0157] Further preferred polypeptide fragments comprising amino acids of the sequences of the EVEN numbered SEQ ID NOs. of the Sequence listing, and polynucleotides encoding the same, are selected from the group consisting of amino acids starting at position one and continuing to any position selected from the group consisting of 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 30 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121,

122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141,
142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161,
162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181,
182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201,
5 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221,
222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241,
242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261,
262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281,
282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301,
10 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321,
322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341,
342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361,
362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381,
382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401,
15 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421,
422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441,
442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461,
462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481,
482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501,
20 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521,
522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541,
542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561,
562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581,
582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601,
25 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621,
622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641,
642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661,
662, 663, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681,
682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 700, 701,
30 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721,
722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738, 739, 740, 741,
742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754, 755, 756, 757, 758, 759, 760, 761,
762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781,

782, 783, 784, 785, and 786, wherein the numbering of amino acids comprising any one fragment is consistent with the polypeptide sequence of any one EVEN numbered SEQ ID of the Sequence listing.

[0158] Further preferred polypeptide fragments comprising amino acids of the sequences of the EVEN numbered SEQ ID NOs. of the Sequence listing, and polynucleotides encoding the same, 5 are selected from the group consisting of amino acids ending at the terminal amino acid of the protein (e.g. position 787) and beginning at any position selected from the group consisting of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 10 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 15 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 20 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 25 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 30 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571,

572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591,
592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611,
612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631,
632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648, 649, 650, 651,
5 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 666, 667, 668, 669, 670, 671,
672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691,
692, 693, 694, 695, 696, 697, 698, 699, 700, 701, 702, 703, 704, 705, 706, 707, 708, 709, 710, 711,
712, 713, 714, 715, 716, 717, 718, 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730, 731,
10 732, 733, 734, 735, 736, 737, 738, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751,
752, 753, 754, 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771,
772, 773, 774, 775, 776, 777, 778, 779, 780, 781, and 782, wherein the numbering of amino acids
comprising any one fragment is consistent with the polypeptide sequence of any one EVEN numbered
SEQ ID of the Sequence listing.

[0159] Further preferred polypeptide fragments of the EVEN numbered SEQ ID NOs. of the
15 Sequence listing, and polynucleotides encoding the same, are selected from the group consisting of
fragments comprising any 50 or 100 consecutive amino acids starting from an amino acid position
selected from the group consisting of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20,
21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47,
48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74,
20 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101,
102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121,
122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141,
142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161,
162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181,
25 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201,
202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221,
222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241,
242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261,
262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281,
30 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301,
302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321,
322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341,
342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361,

362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381,
382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401,
402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421,
422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441,
5 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461,
462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481,
482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501,
502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521,
522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541,
10 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561,
562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581,
582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601,
602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621,
622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641,
15 642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661,
662, 663, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681,
682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 700, 701,
702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721,
722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, and 738, wherein the
20 numbering of amino acids comprising any one fragment is consistent with the polypeptide sequence of
any one EVEN numbered SEQ ID of the Sequence listing.

[0160] These specific embodiments, and other polypeptide and polynucleotide fragment embodiments described herein may be modified as being "at least", "equal to", "equal to or less than", "less than", "at least ____ but not greater than ____" or "from ____ to ____". a specified size or specified N-terminal and/or C-terminal positions. It is noted that all ranges used to describe any embodiment of the present invention are inclusive unless specifically set forth otherwise.

[0161] The above species of polypeptide fragments of the present invention may alternatively be described by the formula "a to b"; where "a" equals the N-terminal most amino acid position and "b" equals the C-terminal most amino acid position of the polynucleotide; and further where "a" equals 30 an integer between 1 and the number of amino acids of the polypeptide sequence of the present invention minus 6, and where "b" equals an integer between 7 and the number of amino acids of the polypeptide sequence of the present invention; and where "a" is an integer smaller than "b" by at least 6.

[0162] The above polypeptide fragments of the present invention can be immediately envisaged using the above description and are therefore not individually listed solely for the purpose of not unnecessarily lengthening the specification. Moreover, the above fragments need not have a GENSET biological activity, although polypeptides having these activities are preferred embodiments 5 of the invention, since even inactive fragments are useful, for example, in immunoassays, in epitope mapping, epitope tagging, as vaccines, as molecular weight markers, and to generate antibodies to a particular portion of the polypeptide.

[0163] The present invention also provides for the exclusion in any polypeptide of any one or more of the above-described fragments, e.g., one or more individual fragments specified by N-terminal 10 and C-terminal positions or of any fragments specified by size in amino acid residues as described above.

Functional definition

Domains

[0164] Preferred polynucleotide fragments of the invention comprise domains of polypeptides 15 of the invention. Such domains may eventually comprise linear or structural motifs and signatures including, but not limited to, leucine zippers, helix-turn-helix motifs, post-translational modification sites such as glycosylation sites, ubiquitination sites, alpha helices, and beta sheets, signal sequences encoding signal peptides which direct the secretion of the encoded proteins, sequences implicated in transcription regulation such as homeoboxes, acidic stretches, enzymatic active sites, substrate binding 20 sites, and enzymatic cleavage sites. Such domains may present a particular biological activity such as DNA or RNA-binding, secretion of proteins, transcription regulation, enzymatic activity, substrate binding activity, etc.

[0165] In a preferred embodiment, domains comprise a number of amino acids that is any integer between 6 and 1000. Domains may be synthesized using any methods known to those skilled 25 in the art, including those disclosed herein. Methods for determining the amino acids which make up a domain with a particular biological activity include mutagenesis studies and assays to determine the biological activity to be tested, as well as bioinformatic methods for recognizing domains, motifs, or signatures, e.g., in a database such as Prosite (Hofmann *et al.*, (1999) Nucl. Acids Res. 27:215-219; Bucher and Bairoch (1994) Proceedings 2nd International Conference on Intelligent Systems for 30 Molecular Biology. Altman et al, Eds., pp53-61, AAAIPress, Menlo Park), Pfam (Sonnhammer, *et al.*, (1997) Proteins. 28(3):405-20; Henikoff *et al.*, (2000) Electrophoresis 21(9):1700-6; Bateman *et al.*, (2000) Nucleic Acids Res. 28(1):263-6), Blocks, Print, Prodom, Sbase, Smart, Dali/FSSP, HSSP,

CATH, SCOP, COG. For a review on available databases, see issue 1 of volume 28 of Nucleic Acid Research (2000).

Epitopes and Antibody Fusions:

[0166] A preferred embodiment of the present invention is directed to epitope-bearing polypeptides and epitope-bearing polypeptide fragments. These epitopes may be "antigenic epitopes" or both an "antigenic epitope" and an "immunogenic epitope". An "immunogenic epitope" is defined as a part of a protein that elicits an antibody response *in vivo* when the polypeptide is the immunogen. On the other hand, a region of polypeptide to which an antibody binds is defined as an "antigenic determinant" or "antigenic epitope." The number of immunogenic epitopes of a protein generally is less than the number of antigenic epitopes (see, e.g., Geysen *et al.*, (1984), PNAS 81:3998-4002). It is particularly noted that although a particular epitope may not be immunogenic, it is nonetheless useful since antibodies can be made to both immunogenic and antigenic epitopes. The present epitopes may be linear (i.e., composed of a contiguous sequence of amino acids repeated along the polypeptide chain) or nonlinear (also called "conformational", i.e., composed of amino acids brought into proximity as a result of the folding of the polypeptide chain).

[0167] An epitope can comprise as few as 3 amino acids in a spatial conformation, which is unique to the epitope. Generally an epitope consists of at least 6 such amino acids, and more often at least 8-10 such amino acids. In preferred embodiment, antigenic epitopes comprise a number of amino acids that is any integer between 3 and 50. Fragments which function as epitopes may be produced by any conventional means (see, e.g., Houghten (1985), PNAS 82:5131-5135), also further described in U.S. Patent No. 4,631,21. Methods for determining the amino acids which make up an epitope include x-ray crystallography, 2-dimensional nuclear magnetic resonance, and epitope mapping, e.g., the Pepscan method described by Geysen, *et al.* (1984); WO 84/03564; and WO 84/03506. Nonlinear epitopes are determined by methods such as protein footprinting (U.S. Patent 5,691,448). Another example is the algorithm of Jameson and Wolf, (1988), Comp. Appl. Biosci. 4:181-186. The Jameson-Wolf antigenic analysis, for example, may be performed using the computer program PROTEAN, using default parameters (Version 4.0 Windows, DNASTAR, Inc., 1228 South Park Street Madison, WI). Epitopes may also be identified *in vivo* by testing for an antigenic response using standard methods.

[0168] All fragments of the polypeptides of the present invention, at least 6, 8, 10, 12, 15, 20, 25, 30, 35, 40, 50, 60, 75, 100, 125, 150, 175, 200, 225, 250, 275, or 300 amino acids residues in length, are included in the present invention as being useful as antigenic linear epitopes. Polypeptides

of the present invention that are not specifically described as immunogenic are not considered non-antigenic as they may be antigenic *in vivo*. The epitope-bearing fragments of the present invention preferably comprise 6 to 50 amino acids (i.e. any integer between 6 and 50, inclusive) of a polypeptide of the present invention. Any number of epitope-bearing fragments of the present invention may also 5 be excluded.

[0169] Nonlinear epitopes comprise more than one noncontiguous polypeptide sequence of at least one amino acid each. Such epitopes result from noncontiguous polypeptides brought into proximity by secondary, tertiary, or quaternary structural features. Preferred polypeptides providing nonlinear epitopes are formed by a contiguous surface of natively folded protein and are thus at least 10 10 amino acids in length, further preferably 12, 15, 20, 25, 30, 35, 40, 50, 60, 75, 100, 125, 150, 175, 200, 225, 250, 275, or 300 amino acids of a polypeptide of the present invention. Additionally, nonlinear epitopes may be formed by synthetic peptides that mimic an antigenic site or contiguous surface normally presented on a protein in the native conformation.

[0170] Immunogenic epitopes may be presented together with a carrier protein, such as an 15 albumin, to an animal system (such as rabbit or mouse) or, if it is long enough (at least about 25 amino acids), without a carrier. However, immunogenic epitopes comprising as few as 8 to 10 amino acids have been shown to be sufficient to raise antibodies capable of binding to, at the very least, linear epitopes in a denatured polypeptide (e.g., in Western blotting).

[0171] Epitope-bearing polypeptides of the present invention are used to induce antibodies 20 according to methods well known in the art including, but not limited to, *in vivo* immunization, *in vitro* immunization, and phage display methods (*see, e.g.*, Sutcliffe, *et al.*, *supra*; Wilson, *et al.*, *supra*, and Bittle, *et al.*, *supra*). If *in vivo* immunization is used, animals may be immunized with free peptide; however, anti-peptide antibody titer may be boosted by coupling of the peptide to a macromolecular carrier using standard methods, such as keyhole limpet hemacyanin (KLH) or tetanus toxoid. Animals 25 such as rabbits, rats and mice are immunized with either free or carrier-coupled peptides, for instance, by intraperitoneal and/or intradermal injection of emulsions containing about 100 µgs of peptide or carrier protein and Freund's adjuvant. Several booster injections may be needed, for instance, at intervals of about two weeks, to provide a useful titer of anti-peptide antibody, which can be detected, for example, by ELISA assay using free peptide adsorbed to a solid surface. The titer of anti-peptide 30 antibodies in serum from an immunized animal may be increased by selection of anti-peptide antibodies, for instance, by adsorption to the peptide on a solid support and elution of the selected antibodies according to methods well known in the art.

[0172] As one of skill in the art will appreciate, and discussed above, the polypeptides of the present invention comprising an immunogenic or antigenic epitope can be fused to heterologous polypeptide sequences. For example, the polypeptides of the present invention may be fused with the constant domain of immunoglobulins (IgA, IgE, IgG, IgM), or portions thereof (CH1, CH2, CH3, any combination thereof including both entire domains and portions thereof) resulting in chimeric polypeptides. These fusion proteins facilitate purification, and show an increased half-life *in vivo* (see, e.g., EPA 0,394,827; and Traunecker *et al.*, (1988), *Nature*. 331:84-86; Fountoulakis *et al.*, (1995) *Biochem.* 270:3958-3964). Additional fusion proteins of the invention may be generated through the techniques of gene-shuffling, motif-shuffling, exon-shuffling, or codon-shuffling (collectively referred to as "DNA shuffling"; see, for example, U.S. Patent Nos.: 5,605,793; 5,811,238; 5,834,252; 5,837,458; and Patten, *et al.* (1997), *Curr Opinion Biotechnol.* 8:724-733; Harayama (1998), *Trends Biotechnol.* 16(2): 76-82; Hansson *et al.*, (1999), *J. Mol. Biol.* 287:265-276; and Lorenzo and Blasco (1998) *Biotechniques.* 24(2):308-313). The present invention further encompasses any combination of the polypeptide fragments listed in this section.

15 Antibodies

Definitions

[0173] The present invention further relates to antibodies and T-cell antigen receptors (TCR), which specifically bind the polypeptides, and more specifically, the epitopes of the polypeptides of the present invention. The antibodies of the present invention include IgG (including IgG1, IgG2, IgG3, and IgG4), IgA (including IgA1 and IgA2), IgD, IgE, or IgM, and IgY. The term "antibody" (Ab) refers to a polypeptide or group of polypeptides which are comprised of at least one binding domain, where a binding domain is formed from the folding of variable domains of an antibody molecule to form three-dimensional binding spaces with an internal surface shape and charge distribution complementary to the features of an antigenic determinant of an antigen, which allows an immunological reaction with the antigen. As used herein, the term "antibody" is meant to include whole antibodies, including single-chain whole antibodies, and antigen binding fragments thereof. In a preferred embodiment the antibodies are human antigen binding antibody fragments of the present invention include, but are not limited to, Fab, Fab' F(ab)2 and F(ab')2, Fd, single-chain Fvs (scFv), single-chain antibodies, disulfide-linked Fvs (sdFv) and fragments comprising either a V_L or V_H domain. The antibodies may be from any animal origin including birds and mammals. Preferably, the antibodies are human, murine, rabbit, goat, guinea pig, camel, horse, or chicken.

[0174] Antigen-binding antibody fragments, including single-chain antibodies, may comprise the variable region(s) alone or in combination with the entire or partial of the following: hinge region, CH1, CH2, and CH3 domains. Also included in the invention are any combinations of variable region(s) and hinge region, CH1, CH2, and CH3 domains. The present invention further includes 5 chimeric, humanized, and human monoclonal and polyclonal antibodies, which specifically bind the polypeptides of the present invention. The present invention further includes antibodies that are anti-idiotypic to the antibodies of the present invention.

[0175] The antibodies of the present invention may be monospecific, bispecific, and trispecific or have greater multispecificity. Multispecific antibodies may be specific for different 10 epitopes of a polypeptide of the present invention or may be specific for both a polypeptide of the present invention as well as for heterologous compositions, such as a heterologous polypeptide or solid support material. *See, e.g.,* WO 93/17715; WO 92/08802; WO 91/00360; WO 92/05793; Tutt, *et al.*, (1991), J. Immunol. 147:60-69; US Patents 5,573,920, 4,474,893, 5,601,819, 4,714,681, 4,925,648; Kostelny *et al.*, (1992), J. Immunol. 148:1547-1553.

[0176] Antibodies of the present invention may be described or specified in terms of the 15 epitope(s) or epitope-bearing portion(s) of a polypeptide of the present invention, which are recognized or specifically bound by the antibody. The antibodies may specifically bind a complete protein encoded by a nucleic acid of the present invention, or a fragment thereof. Antibodies which specifically bind any epitope or polypeptide of the present invention may also be excluded as 20 individual species.

[0177] Thus, another embodiment of the present invention is a purified or isolated antibody capable of specifically binding to any of the polypeptides of the present invention. In one aspect of this embodiment, the antibody is capable of binding to a linear epitope-containing polypeptide comprising at least 6 consecutive amino acids, preferably at least 8 to 10 consecutive amino acids, 25 more preferably at least 12, 15, 20, 25, 30, 40, 50, or 100 consecutive amino acids of a polypeptides of the present invention. In another aspect of this embodiment, the antibody is capable of binding to a nonlinear epitope-containing polypeptide comprising 10 amino acids in length, further preferably 12, 15, 20, 25, 30, 35, 40, 50, 60, 75, or 100 amino acids, further preferably, a contiguous surface of the native conformation of a polypeptide of the present application.

[0178] Antibodies of the present invention may also be described or specified in terms of their cross-reactivity. Antibodies that do not specifically bind any other analog, ortholog, or homolog of the polypeptides of the present invention are included. Antibodies that do not bind polypeptides with less

than 95%, less than 90%, less than 85%, less than 80%, less than 75%, less than 70%, less than 65%, less than 60%, less than 55%, and less than 50% identity (as calculated using methods known in the art and described herein, e.g., using FASTDB and the parameters set forth herein) to a polypeptide of the present invention are also included in the present invention. Further included in the present invention are antibodies, which only bind polypeptides encoded by polynucleotides, which hybridize to a polynucleotide of the present invention under stringent hybridization conditions (as described herein). Antibodies of the present invention may also be described or specified in terms of their binding affinity. Preferred binding affinities include those with a dissociation constant or Kd less than 5×10^{-6} M, 10^{-6} M, 5×10^{-7} M, 10^{-7} M, 5×10^{-8} M, 10^{-8} M, 5×10^{-9} M, 10^{-9} M, 5×10^{-10} M, 10^{-10} M, 5×10^{-11} M, 10^{-11} M, 5×10^{-12} M, 10^{-12} M, 5×10^{-13} M, 10^{-13} M, 5×10^{-14} M, 10^{-14} M, 5×10^{-15} M, and 10^{-15} M.

[0179] The invention also concerns a purified or isolated antibody capable of specifically binding to a mutated GENSET protein or to a fragment or variant thereof comprising an epitope of the mutated GENSET protein.

Preparation of antibodies

15 [0180] The antibodies of the present invention may be prepared by any suitable method known in the art. For example, a polypeptide of the present invention or an antigenic fragment thereof can be administered to an animal in order to induce the production of sera containing "polyclonal antibodies". As used herein, the term "monoclonal antibody" is not limited to antibodies produced through hybridoma technology but it rather refers to an antibody that is derived from a single clone,
20 including eukaryotic, prokaryotic, or phage clone, and not the method by which it is produced. Monoclonal antibodies can be prepared using a wide variety of techniques known in the art including the use of hybridoma, recombinant, and phage display technology.

[0181] Hybridoma techniques include those known in the art (see, e.g., Harlow and Lane, 1988) Antibodies A Laboratory Manual. Cold Spring Harbor Laboratory. pp. 53-242; Hammerling 25 (1981), Monoclonal Antibodies and T-Cell Hybridomas, Elsevier, N.Y. 563-681; Kohler and Milstein, (1975) Nature 256:495; Davis, *et al.*, Basic Methods in Molecular Biology, ed., Elsevier Press, NY (1986), Section 21-2.).

[0182] Briefly, a mouse is repetitively inoculated with a few micrograms of the GENSET protein, or a portion thereof, over a period of a few weeks. The mouse is then sacrificed, and the antibody producing cells of the spleen isolated. The spleen cells are fused by means of polyethylene glycol with mouse myeloma cells, and antibody-producing clones are identified (see, e.g., Engvall,

(1980) *Meth. Enzymol.* 70:419). Selected positive clones can be expanded and their monoclonal antibody product harvested for use.

[0183] Further, Fab and F(ab')2 fragments may be produced, for example, from hybridoma-produced antibodies by proteolytic cleavage, using enzymes such as papain (to produce Fab 5 fragments) or pepsin (to produce F(ab')2 fragments

[0184] Polyclonal antiserum containing antibodies to heterogeneous epitopes in the GENSET protein or a portion thereof can be prepared by immunizing suitable non-human animals (e.g., mouse, rat, rabbit, goat, or horse) with the GENSET protein or a portion thereof, which can be unmodified or modified to enhance immunogenicity (see, e.g., Vaitukaitis *et al.*, (1971) *J. Clin. Endocrinol. Metab.* 10 33:988-991; Ouchterlony *et al.*, (1973) Chap. 19 in: *Handbook of Experimental Immunology* D. Wier (ed) Blackwell). The protein or fragment is typically introduced into the non-human mammal in the presence of an appropriate adjuvant (e.g. aluminum hydroxide, RIBI, etc.). Serum from the immunized animal is collected, treated and tested according to known procedures. If the serum contains polyclonal antibodies to undesired epitopes, the polyclonal antibodies can be purified by 15 immunoaffinity chromatography. Affinity of the antisera for the antigen is determined using standard methods.

[0185] Alternatively, antibodies of the present invention can be produced through the application of recombinant DNA technology or through synthetic chemistry using methods known in the art. In phage display methods, for example, functional antibody domains are displayed on the 20 surface of a phage particle, which carries polynucleotide sequences encoding them, and phage with a desired binding property are selected from a repertoire or combinatorial antibody library (e.g. human or murine) by selecting directly with antigen, typically antigen bound or captured to a solid surface or bead (see, e.g., Brinkman *et al.*, (1995) *J. Immunol Methods*, 182:41-50; Ames *et al.*, (1995), *J. Immunol. Meth.*, 184:177-186.; Kettleborough *et al.*, (1994), *Eur. J. Immunol.*, 24:952-958; Persic *et al.*, (1997), *Gene*, 1879-81; Burton *et al.* (1994), *Adv. Immunol.*, 57:191-280; PCT/GB91/01134; 25 WO 90/02809; WO 91/10737; WO 92/01047; WO 92/18619; WO 93/11236; WO 95/15982; WO 95/20401; and US Patents 5,698,426, 5,223,409, 5,403,484, 5,580,717, 5,427,908, 5,750,753, 5,821,047, 5,571,698, 5,427,908, 5,516,637, 5,780,225, 5,658,727 and 5,733,743). After phage selection, the antibody coding regions from the phage can be isolated and used to generate whole 30 antibodies, including human antibodies, or any other desired antigen binding fragment, and expressed in any desired host including mammalian cells, insect cells, plant cells, yeast, and bacteria (see, e.g., WO 92/22324; Mullinax *et al.*, (1992), *BioTechniques*. 12(6):864-869; and Sawai *et al.*, (1995), *AJRI* 34:26-34; and Better *et al.*, (1988), *Science*. 240:1041-1043).

[0186] Further teaching regarding the preparation of single-chain Fvs and antibodies is provided in U.S. Patents 4,946,778 and 5,258,498; Huston *et al.*, (1991), Meth. Enzymol. 203:46_88; Shu, *et al.*, (1993), PNAS 90:7995-7999; and Skerra, *et al.*, (1988), and Science 240:1038-1040. For some uses, including *in vivo* use of antibodies in humans and *in vitro* detection assays, it may be 5 preferable to use chimeric, shuffled, humanized, or human antibodies (*see e.g.*, Morrison, (1985); Oi *et al.*, (1986), BioTechniques 4:214; Gillies *et al.*, (1989), J. Immunol Methods. 125:191-202; and US Patent 5,807,715; EP 0 239 400; WO 91/09967; US Patent 5,530,101; and 5,585,089, EP 0 592 106; EP 0 519 596; Padlan (1991), Molec. Immunol. 28(4/5):489-498; Studnicka *et al.*, (1994), Protein Engineering. 7(6):805-814; Roguska *et al.*, (1994), PNAS 91:969-973, US Patent 5,565,332, US 10 Patents 4,444,887, 4,716,111, 5,545,806, and 5,814,318; WO 98/46645; WO 98/50433; WO 98/24893; WO 96/34096; WO 96/33735; and WO 91/10741).

[0187] Further included in the present invention are antibodies recombinantly fused or chemically conjugated (including both covalent and non-covalent conjugations) to any heterologous molecule, such as a polypeptide of the present invention, another polypeptide, a label useful for 15 detection assays, or an effector molecule such as a drug or toxin (*see, e.g.*, WO 92/08495; WO 91/14438; WO 89/12624; US Patent 5,314,995; and EP 0 396 387). Fused antibodies may also be used to target the polypeptides of the present invention to particular cell types, either *in vitro* or *in vivo*, by fusing or conjugating the polypeptides of the present invention to antibodies specific for particular cell surface receptors, and may also be used in *vitro* immunoassays and purification methods 20 using methods known in the art (*see e.g.*, Harbor, *et al. supra*; WO 93/21232; EP 0 439 095; Naramura *et al.*, (1994), Immunol. Lett. 39:91-99; US Patent 5,474,981; Gillies *et al.*, (1992), PNAS 89:1428-1432; Fell *et al.*, (1991), J. Immunol. 146:2446-2452).

[0188] The present invention further includes compositions comprising the polypeptides of the present invention fused or conjugated to antibody domains other than the variable regions. For 25 example, the polypeptides of the present invention may be fused or conjugated to an antibody Fc region, or portion thereof, and/or the hinge region, CH1 domain, CH2 domain, or CH3 domain or portions thereof, as well as to portions of IgA or IgM. Methods for fusing or conjugating the polypeptides of the present invention to antibody portions are known in the art. *See e.g.*, US Patents 5,336,603, 5,622,929, 5,359,046, 5,349,053, 5,447,851, 5,112,946; EP 0 307 434, EP 0 367 166; WO 30 96/04388, WO 91/06570; Ashkenazi *et al.*, (1991), PNAS 88:10535-10539; Zheng, *et al.* (1995), J. Immunol. 154:5590-5600; and Vil, *et al.* (1992), PNAS 89:11337-11341.

[0189] Non-human animals or mammals, whether wild-type or transgenic, which express a different species of GENSET polypeptide than the one to which antibody binding is desired, and

animals which do not express a GENSET polypeptide (i.e. a GENSET knock out animal as described herein) are particularly useful for preparing antibodies, as such animals will recognize all or most of the exposed regions of a GENSET protein as foreign antigens, and therefore produce antibodies with a wider array of GENSET epitopes. Moreover, smaller polypeptides with only 10 to 30 amino acids
5 may be useful in obtaining specific binding to any one of the GENSET proteins. In addition, the humoral immune system of animals which produce a species of GENSET that resembles the antigenic sequence will preferentially recognize the differences between the animal's native GENSET species and the antigen sequence, and produce antibodies to these unique sites in the antigen sequence. Such a technique will be particularly useful in obtaining antibodies that specifically bind to any one of the
10 GENSET proteins.

[0190] A preferred embodiment of the invention is a method of specifically binding an antibody or antibody fragment to a GENSET polypeptide. This method comprises the step of contacting a GENSET polypeptide-specific antibody or fragment thereof with a GENSET polypeptide under antibody-binding conditions. Further included is a method of specifically binding an antibody or antibody fragment to an epitope, domain, or fragment of a GENSET polypeptide. This method may be used to, for example, detect, purify, or modify the activity of GENSET polypeptides, as discussed herein.

[0191] Antibodies of the invention can be used to assay protein levels in a test sample or biological sample using methods known to those of skill in the art. Antibody-based methods useful for
20 detecting protein include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as glucose oxidase, horseradish peroxidase, and alkaline phosphatase; radioisotopes, such as iodine (125I, 121I), carbon (14C), sulfur (35S), tritium (3H), indium (113In), and technetium (99Tc); luminescent labels, such luminol, isolumino, theromatic acridinium ester, imidazole,
25 acridinium salt, oxalate ester, luciferin, luciferase, and aequorin; and fluorescent labels, such as fluorescein isothiocyanate, rhodamine, phycoerythrin, phycocyanin, allophycocyanin, o-phthaldehyde, and fluorescamine.

Uses of polynucleotides

Uses of polynucleotides as reagents

30 [0192] The polynucleotides of the present invention may be used as reagents in isolation procedures, diagnostic assays, and forensic procedures. For example, sequences from the GENSET polynucleotides of the invention may be detectably labeled and used as probes to isolate other

sequences capable of hybridizing to them. In addition, sequences from the GENSET polynucleotides of the invention may be used to design PCR primers to be used in isolation, diagnostic, or forensic procedures.

To find corresponding genomic DNA sequences

5 [0193] The GENSET cDNAs of the invention may also be used to clone sequences located in the vicinity, preferably upstream of the cDNAs of the invention on the corresponding genomic DNA. Such sequences may be capable of regulating gene expression, including promoter sequences, enhancer sequences, and other sequences which influence transcription or translation levels.

Use of cDNAs or Fragments thereof to Clone Upstream Sequences from Genomic DNA

10 [0194] Sequences derived from polynucleotides of the inventions may be used to isolate the promoters of the corresponding genes using chromosome walking techniques (e.g., using the GenomeWalker™ kit from Clontech). Once the upstream genomic sequences have been cloned and sequenced, prospective promoters and transcription start sites within the upstream sequences may be identified by comparing the sequences upstream of the polynucleotides of the inventions with 15 databases containing known transcription start sites, transcription factor binding sites, or promoter sequences.

20 [0195] In addition, promoters in the upstream sequences may be identified using promoter reporter vectors, e.g., by placing a reporter gene (e.g., secreted alkaline phosphatase, luciferase, beta-galactosidase, or green fluorescent protein) under the control of regulatory active polynucleotide fragments or variants of the GENSET promoter region located upstream of the first exon of the GENSET gene. A large number of suitable promoter reporter vectors are known in the art. The promoters and other regulatory sequences located upstream of the polynucleotides of the inventions may be used to design expression vectors capable of directing the expression of an inserted gene in a desired spatial, temporal, developmental, or quantitative manner.

25 To find similar sequences

[0196] Polynucleotides of the invention may be used to isolate and/or purify nucleic acids similar thereto using any methods well known to those skilled in the art including the techniques based on hybridization or on amplification described in this section. These methods may be used to obtain the genomic DNAs which encode the mRNAs from which the GENSET cDNAs are derived, mRNAs 30 corresponding to GENSET cDNAs, or nucleic acids which are homologous to GENSET cDNAs or fragments thereof, such as variants, species homologues or orthologs.

Hybridization-based methods

[0197] Techniques for identifying cDNA clones in a cDNA library which hybridize to a given probe sequence are disclosed in Sambrook *et al.*, (1989) Molecular Cloning: A Laboratory Manual. (2ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York), and in Hames and Higgins 5 (1985) Nucleic Acid Hybridization: A Practical Approach (Hames and Higgins Ed., IRL Press, Oxford). The same techniques may be used to isolate genomic DNAs.

[0198] A probe comprising at least 10 consecutive nucleotides from a GENSET cDNA or fragment thereof, is labeled using standard methods with a detectable label such as a radioisotope or a fluorescent molecule. The cDNAs or genomic DNAs in the library are transferred to a nitrocellulose 10 or nylon filter and denatured. After blocking of nonspecific sites, the filter is incubated with the labeled probe for an amount of time sufficient to allow binding of the probe to cDNAs or genomic DNAs containing a sequence capable of hybridizing thereto.

[0199] By varying the stringency of the hybridization conditions used to identify cDNAs or genomic DNAs which hybridize to the detectable probe, cDNAs or genomic DNAs having different 15 levels of identity to the probe can be identified and isolated as described below.

Stringent conditions

[0200] "Stringent hybridization conditions" are defined as conditions in which only nucleic acids having a high level of identity to the probe are able to hybridize to said probe. These conditions may be calculated as follows:

20 [0201] For probes between 14 and 70 nucleotides in length the melting temperature (Tm) is calculated using the formula: $Tm=81.5+16.6(\log (Na^+))+0.41(fraction\ G+C)-(600/N)$ where N is the length of the probe.

[0202] If the hybridization is carried out in a solution containing formamide, the melting temperature may be calculated using the equation: $Tm=81.5+16.6(\log (Na^+))+0.41(fraction\ G+C)-(0.63\% \text{ formamide})-(600/N)$ where N is the length of the probe. 25

[0203] Prehybridization may be carried out in 6X SSC, 5X Denhardt's reagent, 0.5% SDS, 100 µg denatured fragmented salmon sperm DNA or 6X SSC, 5X Denhardt's reagent, 0.5% SDS, 100 µg denatured fragmented salmon sperm DNA, 50% formamide (see, e.g., Sambrook *et al.*, 1986).

[0204] Hybridization is conducted according to standard methods. For probes over 200 30 nucleotides in length, the hybridization may be carried out at 15-25°C below the Tm. For shorter probes, such as oligonucleotide probes, the hybridization may be conducted at 15-25°C below the Tm.

Preferably, for hybridizations in 6X SSC, the hybridization is conducted at approximately 68°C. Preferably, for hybridizations in 50% formamide containing solutions, the hybridization is conducted at approximately 42°C.

[0205] Following hybridization, the filter is washed in 2X SSC, 0.1% SDS at room temperature for 15 minutes. The filter is then washed with 0.1X SSC, 0.5% SDS at room temperature for 30 minutes to 1 hour. Thereafter, the solution is washed at the hybridization temperature in 0.1X SSC, 0.5% SDS. A final wash is conducted in 0.1X SSC at room temperature.

[0206] Nucleic acids which have hybridized to the probe are identified by autoradiography or other conventional techniques.

10 Low and moderate conditions

[0207] Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of formamide result in lowered stringency); salt conditions, or temperature. The above procedure may thus be modified to identify nucleic acids having decreasing levels of identity to the probe sequence. For example, the hybridization temperature may be decreased in increments of 5°C from 68°C to 42°C in a hybridization buffer having a sodium concentration of approximately 1M. Following hybridization, the filter may be washed with 2X SSC, 0.5% SDS at the temperature of hybridization. These conditions are considered to be "moderate" conditions above 50°C and "low" conditions below 50°C. Alternatively, the hybridization may be carried out in buffers, such as 6X SSC, containing formamide at a temperature of 42°C. In this case, the concentration of formamide in the hybridization buffer may be reduced in 5% increments from 50% to 0% to identify clones having decreasing levels of identity to the probe. Following hybridization, the filter may be washed with 6X SSC, 0.5% SDS at 50°C. These conditions are considered to be "moderate" conditions above 25% formamide and "low" conditions below 25% formamide. cDNAs or genomic DNAs which have hybridized to the probe are identified by autoradiography or other conventional techniques.

[0208] Note that variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents (e.g. Denhardt's reagent, BLOTO, heparin, denatured salmon sperm DNA, etc.) used to suppress background in hybridization experiments.

PCR-based methods

[0209] In addition to the above described methods, other protocols are available to obtain homologous cDNAs using GENSET cDNA of the present invention or fragment thereof, as described below.

5 [0210] cDNAs may be prepared by obtaining mRNA from the tissue, cell, or organism of interest, e.g., using mRNA preparation procedures utilizing polyA selection procedures or other techniques known to those skilled in the art. A first primer capable of hybridizing to the polyA tail of the mRNA (e.g. an oligo(T) primer) is hybridized to the mRNA, and a reverse transcription reaction is performed to generate a first cDNA strand (see, e.g., *Current Protocols in Molecular Biology*, John Wiley and Sons, Inc. 1997 and Sambrook, *et al.*, 1989). Typically, such oligo(T) primers comprise an additional sequence upstream of the poly(dT) stretch which facilitates subsequent manipulation of the DNA, such as a restriction site-containing sequence.

10 [0211] The first cDNA strand is then hybridized to a second primer containing at least 10 consecutive nucleotides of a polynucleotide of the invention. Often, the second primer used contains sequences located upstream of the translation initiation site. The second primer is extended to generate a second cDNA strand complementary to the first cDNA strand. Alternatively, RT-PCR may be performed as described above using primers from both ends of the cDNA to be obtained. See, e.g., *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc. 1997 and Sambrook, *et al.*, 1989.

Other protocols

20 [0212] Alternatively, other procedures may be used for obtaining homologous cDNAs. In one approach, cDNAs are prepared from mRNA and cloned into double stranded phagemids using standard methods. The cDNA library in the double stranded phagemids is then rendered single stranded, a biotinylated oligonucleotide comprising the sequence of a fragment of a known GENSET cDNA, genomic DNA or fragment thereof is hybridized to the single stranded phagemids. Hybrids 25 between the biotinylated oligonucleotide and phagemids are isolated and the corresponding phagemids are released from the beads and converted into double stranded DNA using a primer specific for the GENSET cDNA or fragment used to design the biotinylated oligonucleotide. Alternatively, protocols such as the Gene Trapper kit (Gibco BRL), may be used.

As a chromosome marker

[0213] GENSET polynucleotides may be mapped to their chromosomal locations using any methods or techniques known to those skilled in the art including radiation hybrid (RH) mapping (See, e.g., Benham *et al.* (1989) *Genomics* 4:509-517 and Cox *et al.*, (1990) *Science* 250:245-250; and Schuler *et al.*, (1996) *Science* 274:540-546), PCR-based mapping, and Fluorescence *in situ* hybridization (FISH) mapping, as described below.

5 Mapping of cDNAs to Human Chromosomes using PCR techniques

[0214] GENSET cDNAs and genomic DNAs may be assigned to human chromosomes using PCR based methodologies. In such approaches, oligonucleotide primer pairs are designed from the cDNA sequence, and PCR is used to screen a series of somatic cell hybrid cell lines containing defined 10 sets of human chromosomes. Only those somatic cell hybrids with chromosomes containing the human gene corresponding to the GENSET cDNA or genomic DNA will yield an amplified fragment, and the single human chromosome present in all cell hybrids that give rise to an amplified fragment is the chromosome containing that GENSET cDNA or genomic DNA (see, e.g., Ledbetter *et al.*, (1990) *Genomics* 6:475-481).

15 Mapping of cDNAs to Chromosomes Using Fluorescence *in situ* Hybridization

[0215] Fluorescence *in situ* hybridization (FISH) allows the GENSET cDNA or genomic DNA to be mapped to a particular location on a given chromosome. The chromosomes to be used for fluorescence *in situ* hybridization techniques may be obtained from a variety of sources including cell cultures, tissues, or whole blood (see, e.g., Cherif *et al.*, (1990) *PNAS* 87:6639-6643).

20 Use of cDNAs to Construct or Expand Chromosome Maps

[0216] Once the GENSET cDNAs or genomic DNAs have been assigned to particular chromosomes, they may be utilized to construct a high resolution map of the chromosomes on which they are located or to identify the chromosomes in a sample.

[0217] Chromosome mapping involves assigning a given unique sequence to a particular 25 chromosome as described above. Once the unique sequence has been mapped to a given chromosome, it is ordered relative to other unique sequences located on the same chromosome (see, e.g., Nagaraja *et al.*, (1997) *Genome Res.* 1997 Mar;7(3):210-22).

Identification of genes associated with hereditary diseases or drug response

[0218] In another embodiment, any particular GENSET cDNA or genomic DNA may be used 30 as a test probe to associate that GENSET cDNA or genomic DNA with a particular phenotypic characteristic.

[0219] In one embodiment, GENSET cDNAs or genomic DNAs are mapped to a particular location on a human chromosome using standard techniques and the location is searched in Mendelian Inheritance in Man (V. McKusick, Mendelian Inheritance in Man; available on line through Johns Hopkins University Welch Medical Library). Often, this search reveals the region of the human 5 chromosome which contains the GENSET cDNA or genomic DNA to be a very gene rich region containing several known genes and several diseases or phenotypes for which genes have not been identified. The gene corresponding to this GENSET cDNA or genomic DNA thus becomes an immediate candidate for each of these genetic diseases.

[0220] Genomic DNA, mRNA or cDNA from patients with these diseases or phenotypes may 10 then be screened, e.g. using PCR primers from the GENSET cDNA or genomic DNA, or by sequencing, for differences in the expression, size, or sequence of the GENSET cDNA or genomic DNAs in patients relative to in disease-free individuals. Any detected difference indicates a role for the GENSET gene in the disease or phenotype.

Uses of polynucleotides in recombinant vectors

15 [0221] The present invention also relates to recombinant vectors including the isolated polynucleotides of the present invention, and to host cells recombinant for a polynucleotide of the invention, such as the above vectors, as well as to methods of making such vectors and host cells and for using them for production of GENSET polypeptides by recombinant techniques.

Recombinant Vectors

20 [0222] The term "vector" is used herein to designate either a circular or a linear DNA or RNA molecule, which is either double-stranded or single-stranded, and which comprises at least one polynucleotide of interest that is sought to be transferred in a cell host or in a unicellular or multicellular host organism.

25 [0223] In one preferred embodiment, the present invention provides expression vectors comprising either a regulatory polynucleotide or a coding nucleic acid of the invention, or both. Expression vectors may be used to express a GENSET polypeptide for purification, as well as for constructing transgenic animals and also for gene therapy (including *in vivo* and *ex vivo* methods). The encoded protein may be transiently expressed in a host organism or cell or stably expressed in the host organism or cell. The encoded protein may have any of the activities described herein. The 30 encoded protein may be a protein which the host organism lacks or, alternatively, the encoded protein may augment the existing levels of the protein in the host organism.

[0224] Some of the elements which can be found in the vectors of the present invention are described in further detail in the following sections.

General features of the expression vectors of the invention

[0225] Typical expression vectors of the present invention comprise a transcriptional unit comprising an assembly of a genetic element or elements having a regulatory role in gene expression, for example a promoter and a structural or coding sequence which is transcribed into mRNA and eventually translated into a polypeptide, said structural or coding sequence being operably linked to the regulatory elements described above, and appropriate transcription initiation and termination sequences. Additional features may include enhancers, a leader sequence enabling extracellular secretion of translated protein by a host cell, origins of replication, selectable markers permitting transformation of the host cell, ribosome binding sites, polyadenylation signals, splice donor and acceptor sites, transcriptional termination sequences, and 5'-flanking non-transcribed sequences.

Regulatory Elements

[0226] The suitable promoter regions used in the expression vectors according to the present invention are chosen taking into account the cell host in which the heterologous gene has to be expressed.

[0227] A suitable promoter may be heterologous with respect to the nucleic acid for which it controls the expression or alternatively can be endogenous to the native polynucleotide containing the coding sequence to be expressed. Promoter regions can be selected from any desired gene using, for example, CAT (chloramphenicol transferase) vectors and more preferably pKK232-8 and pCM7 vectors. Preferred bacterial promoters are the LacI, LacZ, the T3 or T7 bacteriophage RNA polymerase promoters, the gpt, lambda PR, PL and trp promoters (EP 0036776), the polyhedrin promoter, or the p10 protein promoter from baculovirus (Kit Novagen) (Smith *et al.*, (1983) Mol. Cell. Biol. 3:2156-2165; O'Reilly *et al.* (1992), "Baculovirus Expression Vectors: A Laboratory Manual", W. H. Freeman and Co., New York), the lambda PR promoter or also the trc promoter.

[0228] Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-L. Selection of a convenient vector and promoter is well within the level of ordinary skill in the art.

Selectable Markers

[0229] Selectable markers confer an identifiable change to the cell permitting easy identification of cells containing the expression construct. The selectable marker genes for selection of transformed host cells are preferably dihydrofolate reductase or neomycin resistance for eukaryotic cell culture, TRP1 for *S. cerevisiae* or tetracycline, rifampicin or ampicillin resistance in *E. Coli*, or 5 levan saccharase for mycobacteria, this latter marker being a negative selection marker.

Preferred Vectors

[0230] The present recombinant vectors may be any sort of vector, including, but not limited to, YACs (Yeast Artificial Chromosome), BACs (Bacterial Artificial Chromosome), phage, phagemids, cosmids, plasmids, and linear DNA.

10 Bacterial vectors

[0231] As a representative but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and a bacterial origin of replication derived from commercially available plasmids comprising genetic elements of pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia, Uppsala, Sweden), pGEM1 (Promega Biotec, Madison, 15 WI, USA), pQE70, pQE60, pQE-9 (Qiagen), pbs, pD10, phagescript, psiX174, pbluescript SK, pbsks, pNH8A, pNH16A, pNH18A, pNH46A (Stratagene); ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia); pWLNEO, pSV2CAT, pOG44, pXT1, pSG (Stratagene); pSVK3, pBPV, pMSG, pSVL (Pharmacia); pQE-30 (QIAexpress).

Bacteriophage vectors

20 [0232] The P1 bacteriophage vector may contain large inserts ranging from about 80 to about 100 kb. The construction of P1 bacteriophage vectors is well known in the art (see, e.g., Sternberg (1992) Trends Genet. 8:1-16, Sternberg (1994) Mamm. Genome. 5:397-404, Linton *et al.*, (1993) J. Clin. Invest. 92:3029-3037, McCormick *et al.*, (1994) Genet. Anal. Tech. Appl. 11:158-164.

Viral vectors

25 [0233] Any viral vector can be used to carry out the herein-described methods. In one specific embodiment, the vector is derived from an adenovirus, e.g., human adenovirus type 2 or 5 (see, e.g., Feldman *et al.* (1996) Medecine/Sciences, 12:47-55, Ohno *et al.*, (1994) Science 265:781-784, French patent application No. FR-93.05954).

[0234] Adeno-associated viral vectors are also preferred for the herein-described methods.

[0235] Particularly preferred retroviruses for the preparation or construction of retroviral *in vitro* or *in vitro* gene delivery vehicles of the present invention include retroviruses selected from the group consisting of Mink-Cell Focus Inducing Virus, Murine Sarcoma Virus, Reticuloendotheliosis virus and Rous Sarcoma virus. Particularly preferred Murine Leukemia Viruses include the 4070A and the 1504A viruses, Abelson (ATCC No VR-999), Friend (ATCC No VR-245), Gross (ATCC No VR-590), Rauscher (ATCC No VR-998) and Moloney Murine Leukemia Virus (ATCC No VR-190; PCT Application No WO 94/24298). Particularly preferred Rous Sarcoma Viruses include Bryan high titer (ATCC Nos VR-334, VR-657, VR-726, VR-659 and VR-728). Other preferred retroviral vectors are those described in Roth *et al.*, (1996) Nature Medicine, 2(9):985-991, WO 93/25234, WO 94/06920, Roux *et al.* (1989), PNAS 86:9079-9083, Julian *et al.* (1992), J. Gen. Virol. 73:3251-3255, and Neda *et al.* (1991), J. Biol. Chem. 266:14143-14146.

BAC vectors

[0236] The bacterial artificial chromosome (BAC) cloning system (Shizuya *et al.* (1992), PNAS 89:8794-8797), has been developed to stably maintain large fragments of genomic DNA (100-300 kb) in *E. coli*. A preferred BAC vector comprises a pBeloBAC11 vector that has been described by Kim U-J. *et al.* (1996), Genomics 34:213-218. BAC vectors, and the construction thereof, are well known in the art, and any suitable

Baculovirus

[0237] Another specific suitable host vector system is the pVL1392/1393 baculovirus transfer vector (Pharmingen) that is used to transfect the SF9 cell line (ATCC No. CRL 1711) which is derived from *Spodoptera frugiperda*. Other suitable vectors for the expression of the GENSET polypeptide of the present invention in a baculovirus expression system include those described by Chai *et al.* (1993), Biotechnol. Appl. Biochem. 18:259-273; Vlasak, *et al.* (1983), Eur. J. Biochem. 135:123-126, and Lenhard *et al.*, (1996) Gene. 169:187-190.

25 Delivery Of The Recombinant Vectors

[0238] To effect expression of the polynucleotides and polynucleotide constructs of the invention, the constructs must be delivered into a cell. This delivery may be accomplished *in vitro*, as in laboratory procedures for transforming cell lines, or *in vivo* or *ex vivo*, as in the treatment of certain diseases states.

30 [0239] One mechanism is viral infection, where the expression construct is encapsulated in an infectious viral particle (see, U.S. Patent 5,968,821). The expression construct, preferably a

recombinant viral vector as discussed herein, is used to transduce packaging cells, which then produce infectious viral particles including the expression construct. The particles are then used to transduce eukaryotic cells (see, Miller, A. D. (1990) Blood 76:271; U.S. Patent 6,228,844).

[0240] Replication defective retrovirus comprising a GENSET polynucleotide may be
5 packaged into virions, which can then be used to infect a target cell through the use of a helper virus
by standard techniques (see, e.g., Current Protocols in Molecular Biology, Ausubel, F. M. et al. (eds.)
Greene Publishing Associates, (1989)). Any of a large number of retroviruses can be used and are
well known in the art (see, e.g., Eglitis, et al. (1985) Science 230:1395-1398; Danos and Mulligan
10 (1988) PNAS 85:6460-6464; Wilson, et al. (1988) PNAS 85:3014-3018; Armentano, et al. (1990)
PNAS 87:6141-6145; Huber, et al. (1991) PNAS 88:8039-8043; Ferry, et al. (1991) PNAS 88:8377-
8381; Chowdhury, et al. (1991) Science 254:1802-1805; van Beusechem, et al. (1992) PNAS 89:7640-
7644; Kay, et al. (1992) Human Gene Therapy 3:641-647; Dai, et al. (1992) PNAS 89:10892-10895;
Hwu, et al. (1993) J. Immunol. 150:4104-4115).

[0241] Another viral gene delivery system useful in the present invention utilizes adenovirus-
15 derived vectors. The genome of an adenovirus can be manipulated such that it encodes a gene product
of interest, but is inactivated in terms of its ability to replicate in a normal lytic viral life cycle (see,
e.g., Berkner, et al. (1988) BioTechniques 6:616; Rosenfeld, et al. (1991) Science 252:431-434; and
Rosenfeld, et al. (1992) Cell 68:143-155). Any standard adenoviral vector may be used in the present
invention (see, e.g., Jones, et al. (1979) Cell 16:683; Graham, et al. in Methods in Molecular Biology,
20 E. J. Murray, Ed. (Humana, Clifton, N.J., 1991) vol. 7. pp.109-127).

[0242] Yet another viral vector system useful for delivery of polynucleotides is the adeno-
associated virus (AAV). Adeno-associated virus is a naturally occurring defective virus that requires
another virus, such as an adenovirus or a herpes virus, as a helper virus for efficient replication and a
productive life cycle (see Muzyczka, et al., Curr. Top. Micro. Immunol. (1992) 158:97-129). Any
25 standard AAV vector may be used in the present invention (see, e.g., Flotte *et al.*, (1992) Am. J.
Respir. Cell Mol. Biol. 7:349-356; Am. J. Respir. Cell. Mol. Biol. 7:349-356; Samulski et al. (1989) J.
Virol. 63:3822-3828; and McLaughlin et al. (1989) J. Virol. 62:1963-1973, Tratschin, et al. (1985)
Mol. Cell. Biol. 5:3251-3260, Hermonat, et al. (1984) PNAS 81:6466-6470; Tratschin, et al. (1985)
Mol. Cell. Biol. 4:2072-2081; Wondisford, et al. (1988) Mol. Endocrinol. 2:32-39; Tratschin, et al.
30 (1984) J. Virol. 51:611-619; and Flotte, et al. (1993) J. Biol. Chem. 268:3781-3790).

[0243] Other viral vector systems that may have application in gene therapy have been
derived from herpes virus, vaccinia virus, and several RNA viruses. In particular, herpes virus vectors

may provide a unique strategy for persistence of inserted gene expression in cells of the central nervous system and ocular tissue (Pepose, et al. (1994) *Invest Ophthalmol Vis Sci* 35:2662-2666).

[0244] Several non-viral methods for the transfer of polynucleotides into cells, e.g., mammalian cells, *in vivo*, *in vitro*, or *ex vivo*, are also contemplated by the present invention, and include, without being limited to, calcium phosphate precipitation (Graham *et al.*, (1973) *Virol.* 52:456-457; Chen *et al.* (1987) *Mol. Cell. Biol.* 7:2745-2752); DEAE-dextran (Gopal (1985) *Mol. Cell. Biol.*, 5:1188-1190); electroporation (Tur-Kaspa *et al.* (1986) *Mol. Cell. Biol.* 6:716-718; Potter *et al.*, (1984) *PNAS* 81(22):7161-7165); direct microinjection (Harland *et al.*, (1985) *J. Cell. Biol.* 101:1094-1095); DNA-loaded liposomes (Nicolau *et al.*, (1982) *Biochim. Biophys. Acta.* 721:185-190; Fraley *et al.*, (1979) *PNAS* 76:3348-3352); and receptor-mediated transfection. (Wu and Wu (1987), *J. Biol. Chem.* 262:4429-4432; and Wu and Wu (1988), *Biochemistry* 27:887-892).

[0245] One specific embodiment for a method for delivering a protein or peptide to the interior of a cell of a vertebrate *in vivo* or *in vitro* comprises the step of introducing a preparation comprising a physiologically acceptable carrier and a naked polynucleotide operatively coding for the polypeptide of interest into the interstitial space of a tissue comprising the cell, whereby the naked polynucleotide is taken up into the interior of the cell and has a physiological effect (see, e.g., WO 90/11092, WO 95/11307, Tascon *et al.* (1996), *Nature Medicine* 2(8):888-892, and Huygen *et al.*, (1996) *Nature Medicine* 2(8):893-898). Naked polynucleotides of the invention may also be introduced into cells using particle bombardment (biolistic), e.g., DNA-coated microprojectiles accelerated to a high velocity allowing them to pierce cell membranes and enter cells without killing them (Klein *et al.*, (1987) *Nature* 327:70-73).

[0246] Liposomal preparations for use in the present invention include cationic (positively charged), anionic (negatively charged) and neutral preparations. However, cationic liposomes are particularly preferred because a tight charge complex can be formed between the cationic liposome and the polyanionic nucleic acid. Cationic liposomes have been shown to mediate intracellular delivery of plasmid DNA (Felgner, et al., *PNAS* (1987) 84:7413-7416); mRNA (Malone, et al., *PNAS* (1989) 86:6077-6081); and purified transcription factors (Debs *et al.*, *J. Biol. Chem.* (1990) 265:10189-10192), in functional form. Any of a large number of cationic liposomes can be used. N[1-2,3-dioleyloxy)propyll-N,N,N-triethylammonium (DOTMA) liposomes are particularly useful and are available under the trademark Lipofectin, from GIBCO BRL, Grand Island, N.Y. Other commercially available liposomes include transfectace (DDAB/DOPE) and DOTAP/DOPE (Boehringer).

[0247] Similarly, anionic and neutral liposomes are readily available, such as from AvantiPolar Lipids (Birmingham, Ala.), or can be easily prepared using readily available materials. Such materials include phosphatidyl choline, cholesterol, phosphatidyl ethanolamine, dioleoylphosphatidyl choline (DOPC), dioleoylphosphatidyl glycerol (DOPG), dioleoylphosphatidyl 5 ethanolamine (DOPE), among others. These materials can also be mixed with the DOTMA and DOTAP starting materials in appropriate ratios. Methods for making liposomes using these materials are well known in the art, as are various combinations of liposomes.

[0248] The liposomes can comprise multilamellar vesicles (MLVs), small unilamellar vesicles (SUVs), or large unilamellar vesicles (LUVs), with SUVs being preferred. The various liposome-10 nucleic acid complexes are prepared using methods well known in the art (Straubinger, et al., Methods of Immunology (1983), 101:512-527; U.S. Patent 5,965,421). Generally, the ratio of DNA to liposomes will be from about 10: 1 to about 1: 10 (with about 5:1 to about 1:5 or about 3: 1 to about 1: 3 being preferred). Additionally, liposomes may be targeted to specific cell types by embedding a targeting moiety such as a member of a receptor- receptor ligand pair into the lipid envelope of the 15 vesicle (see, e.g., U.S. Patent 6177433, U.S. Patent 6110490, and P.C.T No. WO9704748).

[0249] The amount of vector to be injected to the desired host organism varies according to the site of injection. As an indicative dose, it will be injected between 0.1 and 100 µg of the vector in an animal body, preferably a mammal body, for example a mouse body.

Secretion vectors

20 [0250] Some of the GENSET cDNAs or genomic DNAs of the invention may also be used to construct secretion vectors capable of directing the secretion of the proteins encoded by genes inserted in the vectors. Such secretion vectors may facilitate the purification or enrichment of the proteins encoded by genes inserted therein by reducing the number of background proteins from which the desired protein must be purified or enriched. Exemplary secretion vectors are described below.

25 [0251] The secretion vectors of the present invention comprise a promoter capable of directing gene expression in the host cell, tissue, or organism of interest, a cloning site for inserting a coding sequence, and a signal sequence from a polynucleotide of the invention is operably linked to the promoter such that the mRNA transcribed from the promoter will direct the translation of the signal peptide. The host cell, tissue, or organism may be any cell, tissue, or organism which 30 recognizes the signal peptide encoded by the signal sequence in the GENSET cDNA or genomic DNA. Suitable hosts include mammalian cells, tissues or organisms, avian cells, tissues, or organisms,

insect cells, tissues or organisms, or yeast. The signal sequences may also be inserted into vectors designed for gene therapy.

[0252] The secretion vector may be DNA or RNA and may integrate into the chromosome of the host, be stably maintained as an extrachromosomal replicon in the host, be an artificial 5 chromosome, or be transiently present in the host. Preferably, the secretion vector is maintained in multiple copies in each host cell.

Cell Hosts

[0253] Another object of the invention comprises a host cell that has been transformed or transfected with one of the polynucleotides described herein, and in particular a polynucleotide either 10 comprising a GENSET polypeptide-encoding polynucleotide regulatory sequence or the polynucleotide coding for a GENSET polypeptide. Also included are host cells that are transformed (prokaryotic cells), transfected (eukaryotic cells), or transduced with a recombinant vector such as one of those described above. Preferred host cells used as recipients for the expression vectors of the invention include prokaryotic host cells such as *Escherichia coli* strains (I.E.DH5- α strain), *Bacillus subtilis*, *Salmonella typhimurium*, and strains from species like *Pseudomonas*, *Streptomyces* and 15 *Staphylococcus*, as well as eukaryotic host cells such as HeLa cells, Cv 1 cells, COS cells, Sf-9 cells, C127 cells, 3T3, CHO, human kidney 293, and BHK cells.

[0254] In addition to encompassing host cells containing the vector constructs discussed herein, the invention also encompasses primary, secondary, and immortalized host cells of vertebrate 20 origin, particularly mammalian origin, that have been engineered to delete or replace endogenous genetic material (e.g., coding sequence), and/or to include genetic material (e.g., heterologous polynucleotide sequences) that is operably associated with the polynucleotides of the invention, and which activates, alters, and/or amplifies endogenous polynucleotides (see, e.g., U.S. Patent No. 5,641,670; WO 96/29411; WO 94/12650; Koller, *et al.*, (1989); and Zijlstra, *et al.* (1989); U.S. Patent 25 Nos. 6,054,288; 6,048,729; 6,048,724; 6,048,524; 5,994,127; 5,968,502; 5,965,125; 5,869,239; 5,817,789; 5,783,385; 5,733,761; 5,641,670; 5,580,734; and in WO96/29411, WO 94/12650; and Koller, *et al.*, (1994).

[0255] GENSET gene expression in mammalian cells, preferably human cells, may be rendered defective, or alternatively may be altered by replacing endogenous GENSET polypeptide-30 encoding genes in the genome of an animal cell by a GENSET polypeptide-encoding polynucleotide according to the invention. These genetic alterations may be generated by homologous recombination using previously described specific polynucleotide constructs.

[0256] Mammal zygotes, such as murine zygotes may be used as cell hosts. For example, murine zygotes may undergo microinjection with a purified DNA molecule of interest. In addition, any one of the polynucleotides of the invention may be introduced in an embryonic stem (ES) cell line, preferably a mouse ES cell line. ES cells, and methods of their isolation and maintenance, are well known in the art, and are described, inter alia, in Abbondanzo *et al.*, (1993), *Meth. Enzymol.*, Academic Press, New York, pp 803-823, Robertson, (1987), *Embryo-derived stem cell lines*; In: E.J. Robertson Ed. *Teratocarcinomas and embryonic stem cells: a practical approach*. IRL Press, Oxford, pp. 71, and Pease and William, (1990), *Exp. Cell. Res.* 190: 209-211.

Transgenic Animals

[0257] The terms "transgenic animals" or "host animals" are used herein to designate animals that have their genome genetically and artificially manipulated so as to include one of the nucleic acids according to the invention. The cells affected may be somatic, germ cells, or both. Preferred animals are non-human mammals and include those belonging to a genus selected from *Mus* (e.g. mice), *Rattus* (e.g. rats) and *Oryctogalus* (e.g. rabbits). In one embodiment, the invention encompasses non-human host mammals and animals comprising a recombinant vector of the invention or a GENSET gene disrupted by homologous recombination with a knock out vector. Thus, the present invention also concerns a transgenic animal containing a nucleic acid, a recombinant expression vector, or a recombinant host cell according to the invention. A further object of the invention comprises recombinant host cells obtained from a transgenic animal described herein.

[0258] Such transgenic animals may be good experimental models in order to study the diverse pathologies related to the increase or decrease of the expression of a given GENSET gene. The transgenic animals may also be used to express a desired polypeptide of interest under the control of the regulatory polynucleotides of the GENSET gene, leading to high yields in the synthesis of this protein of interest, or to tissue-specific expression of the gene.

[0259] In one embodiment, transgenic animals of the present invention are produced by inserting a recombinant polynucleotide of the invention into an embryonic or ES stem cell line (see, e.g., Thomas, *et al.* (1987) *Cell* 51:503-512; Mansour *et al.*, (1988) *Nature* 336:348-352), and isolating, cloning and injecting positive cells into blastocysts, which are then inserted into a female host animal and allowed to grow to term. For more details regarding the production of transgenic animals, and specifically transgenic mice, see US Patents Nos 4,873,191; 5,464,764; 5,789,215, Bradley (1987; In: E.J. Robertson (Ed.), *Teratocarcinomas and embryonic stem cells: A practical*

approach. IRL Press, Oxford, pp.113, Wood, *et al.* (1993), PNAS, 90: 4582-4585, Nagy *et al.*, (1993), PNAS 90: 8424-8428.

[0260] In another embodiment, transgenic animals are produced by microinjecting polynucleotides into a fertilized oocyte. Methods for culturing fertilized oocytes to the pre-
5 implantation stage are described, e.g., by Gordon, *et al.* ((1984) Methods in Enzymology, 101, 414); Hogan, *et al.* ((1986) in Manipulating the mouse embryo, A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y (for the mouse embryo)); Hammer, *et al.* ((1985) Nature, 315, 680 (for rabbit and porcine embryos)); Gandolfi, *et al.* ((1987) J. Reprod. Fert. 81, 23-28); Rexroad, *et al.* ((1988) J. Anim. Sci. 66, 947-953) (for ovine embryos)); and Eyestone, *et al.* ((1989) J.
10 Reprod. Fert. 85, 715-720); Camous *et al.* ((1984) J. Reprod. Fert. 72, 779-785); and Heyman, *et al.* ((1987) Theriogenology 27, 5968 (for bovine embryos)). Pre-implantation embryos are then transferred to an appropriate female by standard methods to permit the birth of a transgenic or chimeric animal, depending upon the stage of development when the transgene is introduced.

[0261] In a preferred embodiment of the present invention, transgenic mammals are generated
15 that secrete recombinant GENSET polypeptides in their milk. Preferably, expression in the mammary gland is accomplished by operably linking the polynucleotide encoding the GENSET polypeptide to a mammary gland specific promoter (e.g., from a casein or lactoglobulin gene) and, optionally, other regulatory elements. Promoter and other regulatory sequences, vectors, and other relevant teachings are provided, e.g., by Clark (1998) J Mammary Gland Biol Neoplasia 3:337-50; Jost, *et al.* (1999) Nat.
20 Biotechnol 17:160-4; U.S. Patent Nos. 5,994,616; 6,140,552; 6,013,857; Sohn, *et al.* (1999) DNA Cell Biol. 18:845-52; Kim, *et al.* (1999) J. Biochem. (Japan) 126:320-5; Soulier, *et al.* (1999) Euro. J. Biochem. 260:533-9; Zhang, *et al.* (1997) Chin. J. Biotech. 13:271-6; Rijnkels, *et al.* (1998) Transgen. Res. 7:5-14; Korhonen, *et al.* (1997) Euro. J. Biochem. 245:482-9; Uusi-Oukari, *et al.* (1997) Transgen. Res. 6:75-84; Hitchin, *et al.* (1996) Prot. Expr. Purif. 7:247-52; Platenburg, *et al.* (1994)
25 Transgen. Res. 3:99-108; Heng-Cherl, *et al.* (1993) Animal Biotech. 4:89-107; and Christa, *et al.* (2000) Euro. J. Biochem. 267:1665-71.

[0262] In another embodiment, the polypeptides of the invention can be produced in milk by introducing polynucleotides encoding the polypeptides into somatic cells of the mammary gland *in vivo*, e.g. mammary secreting epithelial cells. For example, plasmid DNA can be infused through the
30 nipple canal, e.g. in association with DEAE-dextran (see, e.g., Hens, *et al.* (2000) Biochim. Biophys. Acta 1523:161-171), in association with a ligand that can lead to receptor-mediated endocytosis of the construct (see, e.g., Sobolev, *et al.* (1998) 273:7928-33), or in a viral vector such as a retroviral vector, e.g. the Gibbon ape leukemia virus (see, e.g., Archer, *et al.* (1994) PNAS 91:6840-6844). In any of

these embodiments, the polynucleotide may be operably linked to a mammary gland specific promoter, as described above, or, alternatively, any strongly expressing promoter such as CMV or MoMLV LTR.

[0263] The polynucleotides used in such embodiments can either encode a full-length 5 GENSET protein or a GENSET fragment. Typically, the encoded polypeptide will include a signal sequence to ensure the secretion of the protein into the milk.

Uses Of Polypeptides Of The Invention

10 Protein of SEQ ID NO:2 (Internal designation Clone 243525_116-119-1-0-G6-F)

[0264] The cDNA of Clone 243525_116-119-1-0-G6-F (SEQ ID NO:1) encodes protein PCNAlt of SEQ ID NO:2, comprising the amino acid sequence: MPSGEFARICRDLSHIGDAVVISCAKDGVKFSASGELGNNGNIKLSQTSNVDKEEEAVTIEMNEPVQLTFALRYLNFFTAKATPLSSTVTLSMSADVPLVVEYKIADMGHLKYYLAPKIEDEEGS.

15 Accordingly, it will be appreciated that all characteristics and uses of polypeptides of SEQ ID NO:2 described throughout the present application also pertain to the polypeptides encoded by the nucleic acids included in Clone 243525_116-119-1-0-G6-F. In addition, it will be appreciated that all characteristics and uses of the polynucleotides of SEQ ID NO:1 described throughout the present application also pertain to the nucleic acids comprising the human cDNA in Clone 243525_116-119-1-
20 0-G6-F. A preferred embodiment of the invention is directed toward the compositions of SEQ ID NO:1, SEQ ID NO:2, and Clone 243525_116-119-1-0-G6-F. Also preferred are polypeptide fragments having a biological activity as described herein and the polynucleotides encoding the fragments.

[0265] The cDNA of SEQ ID NO:1 is a splice variant of the human proliferating cell nuclear 25 antigen (PCNA), encoding a 123 amino-acid protein of SEQ ID NO:2 named PCNAlt. The protein of SEQ ID:2 comprises the PCNA C-terminal domain.

[0266] Mammalian PCNA is so named because of its initial discovery as a cell cycle-dependent antigen (Miyachi et al, 1978). PCNA plays a critical role in several biological processes that appear disparate but have in common a role in DNA metabolism. PCNAlt is a highly conserved 30 essential component of the DNA replication machinery. The subcellular distribution of PCNAlt changes through the cell cycle, localizing to sites of DNA replication in S-phase of cell cycle. PCNAlt

is a nuclear protein required for leading strand DNA synthesis by DNA polymerase delta, which is part of the essential pathway for DNA replication. Consistent with its essential role in DNA replication, the expression of PCNA lt is very low in quiescent and senescent cells but increases after stimulation by serum or by various growth factors. In addition, PCNA lt and polymerase delta also play essential 5 roles in the repair of damaged DNA. PCNA lt levels are elevated in response to DNA damage *in vivo*, the protein relocates to sites of DNA repair following damage. A direct role of PCNA lt in nucleotide excision repair has been demonstrated. Co-ordination of these two roles of PCNA lt is thus essential for the maintenance of genetic integrity. In addition, PCNA lt can interact with a variety of cellular proteins involved in cell cycle regulation and check point control, such as cyclins and cyclin- 10 dependent kinases.

[0267] In one embodiment, the protein of the invention or part thereof is used to quantify *in vitro* the amount of cell proliferation in a biological sample. PCNA lt is thus used in assays and diagnostic kits for the quantification of cell proliferation in tissue samples such as a tissue section or in cell cultures. The method comprises the steps of contacting a biological sample with a detectably 15 labeled compound capable of selectively binding PCNA lt polypeptides or polynucleotides. For example, a polyclonal or monoclonal antibody, a PCNA lt -binding fragment thereof, or a nucleic acid probe may be used. Preferably, antibody used in this embodiment is directed against nonlinear epitopes of PCNA lt . Preferably, the antibody used in this embodiment specifically binds to PCNA lt polypeptides and not to PCNA polypeptides. Detection can be carried out directly or indirectly with 20 known immunohistological or immunofluorescence techniques. Reagents contained in a kit according to the invention can be directly labeled with generally known molecules, including, but not limited to, enzymes such as alkaline phosphatase and peroxidase and fluorescent dyes such as FITC, rhodamine, and Texas-Red. However, labeling can also occur indirectly by using primary antibodies labeled with molecules such as biotin or digoxigenin which are then detected with a secondary reagent.

[0268] In another embodiment, the current invention provides a method of effectively blocking proliferation or inhibiting the growth of a cell *in vivo* or *in vitro*. Preferably, the present invention can be used to stop unrestrained cell proliferation and to eliminate as many tumor cells as possible. More preferably, it is used to inhibit cell proliferation independent of Gadd45 and p21 control. The method will be performed by administering an antisense oligonucleotide directed against 25 PCNA lt into the cell. Such a method comprises the step of introducing a polynucleotide construct comprising an antisense oligonucleotide directed against PCNA lt to a cell. One strategy for delivering antisense oligonucleotides to targeted cells involves encapsulation or incorporation of the therapeutic bioactive molecules in liposomes, such as cationic liposomes. These liposomes are known to provide 30

a shield against nucleotide degradation *in vivo* and can be targeted to specific areas of the body at which point they slowly release their contents. Alternatively, a polynucleotide construct comprising plasmid DNA operably linked to the antisense oligonucleotide could be used. The nucleotide sequences are administered *in vivo* in a suitable buffer or carrier solution known to those of skill in the art.

[0269] Any compound that inhibits or significantly decreases the expression of PCNA_{Alt} polypeptides can also be used to inhibit replication within one or more cells of the sample. Such compounds can be identified by screening for test substances that decrease PCNA_{Alt} expression. This method of screening comprises the steps of contacting a cell with a test substance and comparing 10 PCNA_{Alt} expression in the cell after exposure to the test substance to that of an unexposed control cell. In addition, any compound that inhibits or significantly decreases the activity of PCNA_{Alt} polypeptides can also be used to inhibit replication within one or more cells of the sample. Such compounds can be identified by screening for test substances that decrease PCNA_{Alt} activity comprising the steps of 15 contacting a cell with a test substance and comparing PCNA_{Alt} activity on the cell replication after exposure to the test substance to that of an unexposed control cell. For example, inhibitor compounds can represent molecules which disrupt the interaction between PCNA_{Alt} and their cellular targets involved in cell progression.

[0270] Alternatively, in another embodiment, the present invention provides a method of effectively stimulating DNA replication in cells. The level or activity of PCNA_{Alt} can be increased in 20 cells to stimulate DNA replication, thereby inducing cell proliferation. Preferably, PCNA_{Alt} is used to increase cell proliferation independent of Gadd45 and p21 control. PCNA_{Alt} levels may be increased by introducing PCNA_{Alt} polynucleotides or polypeptides into a cell in an amount sufficient to specifically stimulate DNA replication of one or more cells within the sample. Preferably, such methods can be performed *in vitro* to increase proliferation of cells in culture. For example, it can be 25 used to maintain cell proliferation under serum starvation conditions that is required for some *in vitro* experiments. On the other hand, an increased level of PCNA_{Alt} is also used to decrease cell proliferation specifically dependent on Gadd45 or p21 control. Whatever the case, the level of PCNA_{Alt} can be increased in cells in any of a number of ways. For instance, purified PCNA_{Alt} protein may be introduced to the cells by microinjection or by liposome or micelle-mediated transport. Such 30 liposomal or micellar microcapsule may optionally be combined with a cell type-specific target, such as an antibody or receptor ligand. Alternatively, PCNA_{Alt} polynucleotides may be introduced to a cell by methods common to the art such as transfection, electroporation, or viral transduction. The vectors mentioned above may also be used to introduce PCNA_{Alt} polynucleotides to a cell.

[0271] The present invention also provides animal models generated by modulating the expression or activity of the present protein in one or more tissues of the animal. Preferably, PCNA_{Alt} expression can be selectively activated or inactivated in a particular cell type using a conditional expression system. These animals can be generated with any inducible method of targeting inactivation of PCNA_{Alt}. Such animals are useful for a number of purposes, because they represent an *in vivo* assay method for testing the efficacy of PCNA_{Alt} inactivation as a candidate strategy potentially useful for the treatment of various pathophysiological aspects of diseases specifically related to an excess of cell proliferation including the cancers listed herein. Such models are also extremely useful in the assessment of combinatory therapies, i. e. PCNA_{Alt} inactivation in addition to one or other compounds of interest. Malignancy conditions to be tested can be induced by any method well known in the art such as chemicals, expression of various oncogenes by transgenesis for example, genotoxic agents (UV, gamma irradiation), etc.

[0272] In another embodiment, the current invention is used to diagnose diseases or disorders associated with unrestrained cell proliferation and in particular in various cancers. Aberrant PCNA_{Alt} expression or localization is a significant prognostic indicator, and is preferably used as a prognostic factor for the clinical course of various cancers. More preferably, it can be measured in case of tumors linked to p53 inactivation. In one such embodiment, the present invention can also provide indicators that are suitable to estimate the efficacy or responsiveness to a form of chemotherapy. The method of detection may preferably occur directly on a section of the sample. The sample material can be present for example as a biopsy, a tissue homogenate, a fine-needle aspirate or tumor resection or can be performed by flow cytometry. The method comprises the steps of contacting a tissue sample obtained from an individual suspected of suffering from the disease or condition or at risk of developing the disease or condition, with a detectably labeled compound capable of selectively binding PCNA_{Alt} polypeptides or nucleic acids. For example, a polyclonal or monoclonal antibody or PCNA_{Alt}-binding fragment thereof or a nucleic acid probe may be used. Preferably, the antibodies used in one such embodiment are directed against nonlinear epitopes of PCNA_{Alt}. Preferably, the antibodies specifically bind to PCNA_{Alt} polypeptides and not to PCNA polypeptides. Detection can be carried out directly or indirectly with known immunohistological and immunofluorescent techniques. Reagents contained in a kit according to the invention can be directly labeled with generally known molecules, including, but not limited to enzymes such as alkaline phosphatase and peroxidase and fluorescent dyes such as FITC, rhodamine, and Texas-Red. However, labeling can also occur indirectly by using secondary antibodies labeled with molecules such as biotin, digoxigenin or the like and are then detected with a secondary reagent. PCNA_{Alt} immunoreactivity can be regarded as indicative of

neoplasia but preferably with at least one other marker. The second marker index to be detected is selected from the group comprising a marker for transformed cells, protein being overexpressed in a neoplastic cell, or a protein expressed in a form normally not present in the cell. For example, the procedure can be adapted for double immunostaining of PCNA_{Alt} together with alpha-actin, 5 bromodeoxyuridine, keratin, type IV collagen and vimentin. In addition, the protein representing the second marker may preferably represent a tissue specific marker. Thus, a kit may contain an antibody directed against PCNA_{Alt} and antibodies against one or more of the markers mentioned above. The condition of an individual can be monitored continuously and the quantified amount of this particular protein measured in the pathological sample can be compared with the amount quantified in a 10 biological sample of a normal individual or with previous samples from the same individual. The determination of all indices mentioned above can occur at the same time, sequentially or after one another.

[0273] In one other embodiment, the present invention is useful in diagnosing patients with an excess amount of PCNA_{Alt} and monitor PCNA_{Alt} expression in such conditions. Preferably, the present 15 invention provides a method to diagnose and the monitor diseases such as systemic lupus erythematosis disease and malignant lymphomas. The method comprises the steps of contacting a tissue sample obtained from an individual suspected of suffering from the disease or condition or at risk of developing such a disease or condition, with a detectably labeled compound capable of selectively binding PCNA_{Alt} polypeptides or nucleic acids. For example, a polyclonal or monoclonal 20 antibody or a PCNA_{Alt}-binding fragment thereof or a nucleic acid probe may be used. Preferably, the antibodies used in one such embodiment are directed against nonlinear epitopes of PCNA_{Alt}. Preferably, the antibodies specifically bind to PCNA_{Alt} polypeptides and not to PCNA polypeptides. Thus, the amount of PCNA_{Alt} can be monitored and quantified in a subject in the pathological sample and be compared with the amount quantified in a biological sample of a normal individual.

25 [0274] A further embodiment of the present invention is to provide novel methods and compositions useful for the treatment or prevention of diseases and conditions related to excessive cell proliferation and preferably with malignancy conditions. Such methods comprise the administration of a therapeutically-effective amount of antisense PCNA_{Alt} oligonucleotides to mammals suffering from the disease or condition, where "effective amount" is meant a concentration of antisense 30 oligonucleotides capable of significantly reducing cell proliferation. The compositions of the invention are preferably delivered to an individual in a pharmaceutically acceptable carrier, such as a saline solution or other physiological buffer suitable for administration to a patient. The particular amount of the compositions of the invention that will be administered to the mammal for any

particular condition will depend on the clinical condition of the patient, and other factors such as the weight, age, and route of delivery. Such compositions can be administered by any suitable route. For treatment purposes, the composition can also be incorporate into liposomes, such as cationic liposomes to deliver the therapeutic bioactive molecules to targeted cells. These compositions can comprise 5 antisense oligonucleotides directed against PCNA_{Alt}, and, optionally, one or more other compounds of interest. This co-administration may be by simultaneous administration or by separate or sequential administrations. All of these additional components may be either obtained from natural sources or produced by recombinant genetic engineering techniques and/or chemical modification.

Protein of SEQ ID NO:4 (Internal designation Clone 643144_181-17-2-0-A12-F)

10 [0275] The cDNA of Clone 643144_181-17-2-0-A12-F (SEQ ID NO:3) encodes Lectir of SEQ ID NO:4, comprising the amino acid sequence: MQDEDGYITLNKTRKPALVSGSASSSWRVMALILLILCVGMVGLVALGIWSVMQRNY LQDENENRTGTLQQQLAKRFCQYVVVKQSELKGTFKGHKCSPCDTNWRYYGDSCYGFFRHNLT WEESKQYCTDMNATLLKIDNRNIVEYIKARTHLLIRWVGLSRQKSNEVWKWEDGSVISENMFE 15 FLEDGKGNMNCAVFHNGKMHPTFCENKHYLMCERKAGMTKVDQLP. Accordingly, it will be appreciated that all characteristics and uses of the polypeptides of SEQ ID NO:4 described throughout the present application also pertain to the polypeptides encoded by the nucleic acids comprising the human cDNA in Clone 643144_181-17-2-0-A12-F. In addition, it will be appreciated that all characteristics and uses of the polynucleotides of SEQ ID NO:3 described throughout the present 20 application also pertain to the nucleic acids included in Clone 643144_181-17-2-0-A12-F. A preferred embodiment of the invention is directed toward the compositions of SEQ ID NO:4, SEQ ID NO:3, and Clone 643144_181-17-2-0-A12-F.

[0276] Lectir is a polymorphic variant of the C-type Lectin-like receptor 2 (GenBank accession number AAF36777). Lectir possesses an anchor signal, a transmembrane domain 25 (VMALILLILCVGMVGLVALGIW), a lectin C-type domain (QYCTDMNATLLKIDNRNIVEYIKARTHLLIRWVGLSRQKSNEVWKWEDGSVISENMFEFLED GKGNMNCAVFHNGKMHPTFCENKHYLMCE) and an extracellular link domain (RHNLTWEESKQYCTDMNATL). Lectir is homologous to the oxidized LDL receptor (LOX-1). Lectir is expressed in myeloid cells such as monocytes, dendritic cells and granulocytes.

30 [0277] Lectir is involved in antigen capture and phagocytosis of apoptotic bodies. Phagocytosis of aged and apoptotic cells is an essential process to protect normal healthy cells from the harmful contents and debris of dying cells. Apoptosis is accompanied by the clustering of

intracellular autoantigens, which are selectively-cleaved and phosphorylated, and by the exposure of anionic phospholipids such as phosphatidyl-serine. Modified lipoproteins (e.g., oxidized lipoproteins) play a crucial role in apoptosis signaling. Lectins and C-type lectin like receptors are expressed both by cells undergoing apoptosis and by phagocytic cells. Lectir is a phagocytic receptor that binds to 5 phospholipids, captures apoptotic bodies and directs them from the extracellular space to a specialized processing compartment of myeloid cells.

[0278] An embodiment of the invention is directed to compositions comprising an antibody directed against a Lectir polypeptide or against a Lectir polypeptide fragment. Preferably, the antibody specifically binds to the Lectir polypeptide and not to the C-type Lectin-like receptor-2 10 polypeptide. An additional preferred embodiment of the invention is a method of binding Leclir polypeptides with a Leclir-specific antibody or Leclir-binding fragment thereof. This method comprises the step of contacting a Leclir polypeptide with a Leclir-binding antibody or Leclir-binding fragment thereof under conditions that allow binding. This method may be applied to detection and purification of Leclir, as well as targeting Leclir-expressing cells. These aspects are discussed in detail 15 herein.

[0279] A preferred aspect of the invention is a host cell recombinant for polynucleotides encoding a Leclir polypeptide or a biologically active fragment thereof. An additional preferred aspect is a host cell recombinant for polynucleotides capable of directing Leclir expression. Preferably, the polynucleotides capable of directing Leclir expression are located in the 5' regulatory region of the 20 Leclir gene. Further preferably, these polynucleotides are located within 500 base pairs of the Leclir coding region. These polynucleotides preferably comprise a promoter sequence.

[0280] An embodiment of the present invention includes methods of purifying Lectir polypeptides. This method comprises the steps of: i) obtaining a cell capable of expressing a Lectir polypeptide; ii) growing said cell under conditions suitable to produce said polypeptide; and iii) 25 purifying said polypeptide. An example of this method comprises the steps of: i) transfecting a mammalian host cell with a recombinant expression vector comprising a polynucleotide of the present invention, and ii) purifying the produced protein. Leclir purification of the protein can be done following any technique well-known to those skilled in the art. Preferably, a Leclir-binding antibody or an antigen binding fragment thereof may be bound to a chromatographic support to form an affinity 30 chromatography column. Preferably, the antibodies are those described above.

[0281] An embodiment of the present invention relates to methods of using the polypeptides and the polynucleotides of the present invention to enhance phagocytosis of accumulating apoptotic

cells. Any compositions and methods containing, e.g., Lectir polypeptide or part thereof, a polynucleotide encoding the protein, or a compound that increases the expression or activity of Lectir can be used.

[0282] In a preferred embodiment, a Lectir polypeptide is used in a method to enhance 5 phagocytosis of apoptotic cells. This method comprises the step of introducing Lectir polypeptide to a cell. Preferred cells are phagocytic cells, such as dendritic cells and macrophages. Lectir polypeptides may be introduced by introducing a polynucleotide construct comprising polynucleotides that encode 10 Lectir polypeptides or a biologically active fragment thereof to a cell. This method may be directed to the removal of apoptotic cells from a sample of cells to be cultured. Additionally, this method may be applied to reduce the severity of diseases that result from cellular accumulation such as rheumatoid 15 arthritis, systemic lupus erythematosus, psoriatic arthritis, asthma, bronchitis, sarcoidosis, and pulmonary fibrosis. This method may also be applied when treating a patient with apoptosis-inducing drugs such as, e.g., cisplatin and carboplatin, for reducing cellular accumulation of apoptotic cells. Notably, apoptosis-inducing drugs are widely used for treating tumor proliferation.

15 [0283] In another embodiment, the methods of the present invention relate to the administration of a recombinant expression vector comprising one of the polynucleotides of the invention to a patient suffering from a disease associated with accumulation of apoptotic cells. Preferred expression vectors include viral vectors, especially adenoviral and lentiviral vectors.

[0284] In still another embodiment, genetic modification of a cell with a vector comprising 20 one of the polynucleotides of the invention may be accomplished using one or more techniques well known in the gene therapy field. For example, one of the methods described in Mulligan (Mulligan, Science, 260:926-32 (1993)), which disclosure is hereby incorporated by reference in its entirety, can be used.

[0285] In still another embodiment, the compositions of the present invention comprise a 25 substance that increases Lectir expression. Additionally, the methods of the present invention relate to methods of screening test substances that increase Lectir expression. These methods comprise the steps of: i) contacting a cell with a test substance; and ii) comparing Lectir expression in the cell after exposure to the test substance to that of an unexposed control cell. Preferably, the test substance modifies the expression of Lectir in monocytes, dendritic cells and/or granulocytes while not in other 30 cell types.

[0286] Another embodiment of the present invention is directed to compositions comprising substances that increase Lectir phagocytic activity. Additionally, the methods of the present invention

relate to methods of screening test substances that increase Lectir expression. These methods comprise the steps of: i) contacting a cell with a test substance; and ii) comparing Lectir activity in the cell after exposure to the test substance to that of an unexposed control cell. Preferably, the test substance modifies Lectir activity in monocytes, dendritic cells and/or granulocytes while not in other 5 cell types. Several methods, including but not limited to the method described by Simpson et al. (J Immunol Methods 29:221-6 (1979)), which disclosure is hereby incorporated by reference in its entirety, are available for quantifying phagocytic activity. Such a method can also be used for verifying *in vitro* effectiveness of compositions and methods of the present invention to enhance phagocytosis.

10 [0287] In a preferred embodiment, Lectir polypeptides or a substance that increases Lectir expression or Lectir activity can be processed in accordance with conventional methods of pharmacy to produce medicinal agents for administration to patients. Thus, the pharmaceutical composition comprising Lectir or part thereof or a substance that increases Lectir expression may be made up in a solid form (e.g. granules for oral administration, powders for inhalation) or in a liquid form (e.g. 15 solutions for oral administration or for injection). Effective doses of the compositions of the present invention for treating a patient suffering from disorders associated with accumulation of apoptotic cells can be determined according to the relevant techniques.

[0288] Another embodiment of the present invention relates to methods of using an antibody directed against Lectir polypeptides for purifying myeloid cells. Preferred antibodies are those 20 described above. Such a method for myeloid cells comprises the steps of: i) labeling an Lectir-specific antibody with a molecule that can be used to provide a detectable signal, ii) running a body fluid through a sorting apparatus (e.g., a fluorescence-activated cell sorter or a magnetic activated cell-sorting apparatus, containing the labeled anti-Lectir antibody). Myeloid cells, which include monocytes, granulocytes, and megakaryocytes monitor for the presence of foreign bodies, provide 25 protection against neoplastic cells, scavenge foreign materials, and produce platelets. A highly purified population of myeloid cells is necessary for a variety of *in vitro* experiments and *in vivo* indications. For example, myeloid cells find use in i) enriching the hematopoietic system of a host deficient in any class of myeloid cells; and ii) detecting diseases associated with myeloid cell dysfunction.

30 Protein of SEQ ID NO:6 (Internal designation Clone 212950.cREC_116-075-2-0-H1-F)

[0289] The cDNA of Clone 212950.cREC_116-075-2-0-H1-F (SEQ ID NO:5) encodes v1ADAM20 of SEQ ID NO:6, comprising the amino acid sequence:

MVQLHQDTDPQIPKGQPCTLNSSEGGARPAVPHTLFSSALDRWLHNDSFIMAVGEPLVHIRVT
LLLLWFGMFLSISGHSQARPSQYFTSPEVVIPLKVISRGRAKAPGWLSYSLRFGGQRYTIVHM
RVNKLLFAAHLPVFTYTEQHALLQDQPFIQDDCYYHGYVEGPESLVALSTCSGGFLGMLQI
NDLVYEIKPISVSATFEHLVYKIDSDDTQFPPMRCGLTEEKIAHQMLEQLSYNFTLKQSSFVGW
5 WTHQRFVELVVVDNIRYLFSQSNATTVQHEVFNVVNIVDSFYHPLEVDVILTGIDIWTASNP
LPTSGDLDNVLEDFSIWKNYNLNNRLQHDVAHLFIKTQGMKLGVAYVKGICQNPFTGVD
VFEDNRLVVFATLGHELGHNLGMQHDTQWCVCSELQWCIMHAYRKVTTKFSNCSYAQYWD
STISSGLCIQPPPYPGNIFRLKYCGNLVVEEGEECDCGTIRQCACDPCCLLNCTLHPGAACAFGI
CCKDCFKLPSGTLCRQQVGECDLPEWCNGTSHQCPDDVYVQDGISCNVNAFCYEKTCNNHDI
10 QCKEIFGQDARSASQSCYQEINTQGNRGHCGIVGTTYVKCWTPDIMCGRVQCENVGVIPNLI
EHSTVQQFHLNDTCWGTDYHLGMAIPDICEVKDGTVCPEKICIRKKCASMVHLSQACQPK
TCNMRGICNNKQHCHCNHEWAPPYCKDKGYGGSADSGPPPKNMEGLNVMGKLRYLSLLC
LLPLVAFLLFCLHVLFKKRTKSKEDEEG. Accordingly, it will be appreciated that all
characteristics and uses of the polypeptides of SEQ ID NO:6 described throughout the present
15 application also pertain to the polypeptides encoded by the nucleic acids comprising the human cDNA
in Clone 212950.cREC_116-075-2-0-H1-F. In addition, it will be appreciated that all characteristics
and uses of the polynucleotides of SEQ ID NO:5 described throughout the present application also
pertain to the nucleic acids included in Clone 212950.cREC_116-075-2-0-H1-F. A preferred
embodiment of the invention is directed toward the compositions of SEQ ID NO:6, SEQ ID NO:5, and
20 Clone 212950.cREC_116-075-2-0-H1-F.

[0290] Preferred v1ADAM20 polypeptides for uses in the methods described below include
the polypeptides comprising the amino sequence of:
MVQLHQDTDPQIPKGQPCTLNSSEGGARPAVPHTLFSSALDRWLHNDSFI.

[0291] Also preferred are the polypeptide fragments having a biological activity of
25 proteolytically processing sperm surface proteins and the polynucleotides encoding the fragments.

[0292] v1ADAM20 is a sperm protein that is a member of the ADAM family, and that
displays all structural elements that are characteristic of this family. Notably, v1ADAM20 displays a
catalytic site of an active Zn²⁺- metalloprotease (HELGHNLGMQHD), one disintegrin consensus
sequence (EEGEECDCG) and one transmembrane domain. The ADAM family of proteins contains a
30 disintegrin and metalloproteinase domain. Members of this family are cell surface proteins with a
unique structure possessing both potential adhesion and protease domains. Although all ADAMs have
the same domain organization, members of the ADAM do not all display the same functions. The
functions displayed by members of the ADAM family are regulated by processes that include

alternative splicing, differential gene expression, dimerization, and proteolytic processing. ADAMs are involved in diverse biological processes such as spermatogenesis, fertilization, myoblast fusion and neuron proliferation (Wolfsberg *et al*, Dev Biol. 180:389-401 (1996)).

[0293] *vlADAM20* is a novel membrane-anchored sperm surface protease that is involved in spermatogenesis. *vlADAM20* is involved in remodeling the developing sperm surface, *i.e.*, proteolytically processing other sperm surface proteins.

[0294] An embodiment of the invention is directed to a composition comprising a *vlADAM20* polypeptide sequence of SEQ ID NO:6.

[0295] A further embodiment of the invention is directed to a composition comprising a *vlADAM20* polypeptide fragment having biological activity of proteolytically processing sperm surface proteins.

[0296] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence of SEQ ID NO:5 encoding a *vlADAM20* polypeptide.

[0297] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence encoding a *vlADAM20* polypeptide fragment having biological activity of proteolytically processing sperm surface proteins.

[0298] A further embodiment of the invention is directed to an antibody or an antigen-binding fragment that binds to a *vlADAM20* polypeptide or against a *vlADAM20* polypeptide fragment. Preferably, the antibody specifically binds to the *vlADAM20* polypeptide and not to the *ADAM20* polypeptide. Preferably, the antibody or antigen-binding fragment recognizes an epitope comprising one or more of the 50 amino-acids located at the amino-terminal extremity of *vlADAM20*, wherein one or more of these amino-acids are required for antibody binding. Even more preferably, the antibody recognizes the QLHQD TD P QIPKGQ PCT amino-acid sequence or the LNSSEGGAR amino-acid sequence. As used herein, the term *anti-vlADAM20 antibody* includes both intact molecules as well as active fragments thereof, such as those capable of binding antigens.

[0299] The present invention also relates to a method of binding an anti- *vlADAM20* antibody to a *vlADAM20* polypeptide comprising the step of: contacting a *vlADAM20* polypeptide with said antibody or antigen-binding fragment thereof under conditions that allow binding to take place. Such conditions are well known to those skilled in the art. Additionally, the present invention also relates to a method of binding an anti- *vlADAM20* antibody or antigen-binding fragment thereof to a cell expressing a *vlADAM20* polypeptide or fragment thereof comprising the step of: contacting a

cell expressing a vlADAM20 polypeptide or part thereof with said antibody or antigen-binding fragment under conditions that allow binding to take place. Preferably, the antibody or fragment used in these methods is vlADAM20-specific, as discussed above. Such methods are for example useful for purifying testis cells as further described herein.

5 [0300] Another embodiment of the present invention relates to methods of using compositions comprising an antibody directed against vlADAM20 polypeptides for purifying testis cells. Preferably, such compositions comprise the preferred antibodies described above. Such a method for purifying testis cells comprises the steps of: i) labeling by standard methods of the anti- vlADAM20 antibody with a molecule that can be used to provide a detectable signal, ii) running a biopsied testis
10 sample or mammalian cells through a sorting apparatus, e.g. a fluorescence-activated cell sorter or a magnetic activated cell-sorting apparatus, containing the labeled anti- vlADAM20 antibody. Purifying testis cells are useful for *in vitro* analysis and diagnosis of various diseases associated with testis including, but not limited to, testicular cancer, testicular intraepithelial neoplasia, testicular microlithiasis and sterility.

15 [0301] An embodiment of the present invention relates to a method of producing vlADAM20 polypeptides comprising the steps of: i) transfecting a host cell with a recombinant expression vector comprising a polynucleotide of the present invention; and optionally ii) purifying the produced protein. The purification of the protein can be done following any technique well known to those skilled in the art. Preferably, an antibody directed against vlADAM20 or part thereof may be bound to a
20 chromatographic support to form an affinity chromatography column. Even more preferably, the antibody recognizes the 50 amino-terminal amino acids of vlADAM20. Alternatively, vlADAM20 polypeptides may be produced by a method comprising the step of: i) transfecting a host cell with a polynucleotide capable of directing vlADAM20 expression. Preferably, the polynucleotides capable of directing vlADAM20 expression are located in the 5' regulatory region of the vlADAM20 gene.
25 Further preferably, these polynucleotides are located within 500 base pairs of the vlADAM20 coding region. These polynucleotides preferably comprise a promoter sequence. Techniques known in the art for introducing polynucleotide sequences to endogenous sequences are described in U.S. Patent 5641670 and PCT WO9629411, which disclosures are hereby incorporated by reference in their entireties.

30 [0302] In an embodiment, a vlADAM20 polypeptide or a biologically active fragment thereof can be used in a “cocktail” of proteases that is able to digest a wide range of proteins. Such protease cocktails are useful in laboratory assays to degrade undesirable proteins in a sample, for example for removing proteins in a DNA preparation or for removing enzymes after any enzymatic reaction.

[0303] In another embodiment, vlADAM20 or a biologically active fragment thereof can be used in combination with a detergent for the removal of stains having a protein component, similar to the use of proteases described in U.S. Patent 5,599,400, which disclosure is incorporated by reference in its entirety. The composition can contain known detergent constituents such as, e.g. surfactants, 5 foam enhancers and fillers. The detergent preferably contains between 0.001% to 10% vlADAM20 polypeptides. vlADAM20 polypeptides can be included in a detergent composition or can be combined with other constituents at the time of use as an additive. The detergent additive can be formulated as a liquid, powder, granulate or slurry.

[0304] Still another embodiment relates to compositions and methods for using vlADAM20 polypeptides as a contraceptive immunogen. It was shown that men and women who spontaneously produce anti-sperm antibodies are infertile but otherwise healthy (Bronson *et al*, Fertil Steril. 42:171-83 (1984)). Such a method comprises the step of: administering to an individual a vlADAM20 polypeptide as described in U.S. Patent 6,197,940, which disclosure is hereby incorporated in its entirety, in an effective amount to reduce the fertility of that individual *via* generation of antibodies to 15 vlADAM20. Preferably, immunogenic vlADAM20 polypeptides are used in such compositions. The administered amount of immunogenic vlADAM20 polypeptides depends upon factors such as route of administration, species, and the use of booster administration. For example, a dose of about 0.1 to 100 micrograms of vlADAM20 polypeptides per kg of body weight may be used. Compositions of the present invention may be prepared as both human and veterinary vaccine formulations. Contraceptive 20 vaccines can be produced by combining vlADAM20 polypeptides with a pharmaceutically suitable carrier and with an adjuvant that contains non-specific stimulators of the immune system such as, e.g., immunogenic fragments of *Bordatella pertussis*.

[0305] Compositions of the present invention may be made up in a solid form (e.g. granules for oral administration) or in a liquid form (e.g. solutions for oral administration or for injection). 25 Compositions of the present invention are included in the carrier in an amount that is effective to reduce the fertility of the subject being treated. Compositions of the present invention may also comprise other contraceptive molecules. For example, a female contraceptive vaccine comprising vlADAM20 polypeptides may also comprise PH30 beta chain proteins that are described in U.S. Patent 5,693,496, which disclosure is incorporated hereby by reference in its entirety, and a male 30 contraceptive vaccine may also comprise SP-22 polypeptides that are described in U.S. Patent 6,197,940, which disclosure is incorporated hereby by reference in its entirety.

[0306] A further embodiment of the present invention is directed to substances that decrease vlADAM20 expression, and to a method of screening for such substances comprising the steps of: i)

contacting a cell with a test substance, ii) comparing vLADAM20 expression in an exposed cell to that of an unexposed control cell, iii) quantifying said expression levels, and iv) determining the ratio of vLADAM20 expression in an exposed cell relative to the expression in an unexposed cell. Preferably, vLADAM20 expression is studied in testis cells. Preferably, compositions comprising substances that 5 decrease vLADAM20 expression can be administered to male individuals for contraceptive purposes.

[0307] A further embodiment of the present invention is directed to substances that decrease vLADAM20 activity, and to a method of screening for such substances comprising the steps of: i) contacting a cell with a test substance, ii) comparing vLADAM20 proteolytic activity in an exposed cell to that of an unexposed control cell, iii) quantifying said expression levels, and iv) determining the 10 ratio of vLADAM20 proteolytic activity in an exposed cell relative to the expression in an unexposed cell. Preferably, vLADAM20 expression is studied in testis cells. vLADAM20 proteolytic activity can for example be measured by the alpha 2M complex formation assay as described by Frosty Loechel *et al* (J Biol Chem 273:16993-7 (1998)). Preferably, compositions comprising substances that decrease vLADAM20 expression can be administered to male individuals for contraceptive purposes.

15 [0308] As used herein, the term *antagonist* includes substances that decrease vLADAM20 expression, substances that decrease vLADAM20 activity, and antibodies or antigen-binding fragments that inhibit the activity of vLADAM20.

[0309] Another embodiment of the present invention relates to methods of using compositions comprising a vLADAM20 antagonist for blocking vLADAM20 proteolytic activity. Preferably, the 20 antagonist is one of the preferred antibodies described above. Such compositions can be administered to a cell, a tissue sample or an individual. Preferably, such compositions are administered to male individuals for blocking spermatogenesis in contraceptive purposes.

[0310] Still another embodiment of the present invention relates to methods of using compositions comprising an anti- vLADAM20 antibody for agglutinating and immobilizing sperm. 25 Preferably, such compositions comprise the preferred antibodies described above. Such compositions can be administered to, e.g., a cell, a tissue sample or a patient. Preferably, such compositions are administered to female individuals for blocking fertilization in contraceptive purposes.

Protein of SEQ ID NO:8 (Internal designation Clone 1000849866_181-44-3-0-A9-F)

[0311] The cDNA of SEQ ID NO:7 of clone 1000849866_181-44-3-0-A9-F encodes 30 Lipoglobulin of SEQ ID NO:8, comprising the sequence:
SLGALLLLSACLA VSAGPVPTPPDNIQVQENFNISRIY GKWYNLAIGSTCPWLKKIMDRMTV
STLVLGEGATEAEISMTSTRWRKGVCEETSGAYEKTDGKFLYHKSKWNITMESYVVHTN

YDEYAIFLT KFKSRHHGPTITAKLYGRAPQLRETLLQDFRVVAQGVGIPEDSIFTMADRGECPV
GEQEPEPILIPRVRRAATPRRGRIRGWATGN. Accordingly, it will be appreciated that all
characteristics and uses of the polypeptide of SEQ ID NO:8 described throughout the present
application also pertains to the polypeptide encoded by the nucleic acids comprising the human cDNA
5 in clone 1000849866_181-44-3-0-A9-F. In addition, it will be appreciated that all characteristics and
uses of the polynucleotides of SEQ ID NO:7 described throughout the present application also pertains
to the polypeptide encoded by the nucleic acids included in clone 1000849866_181-44-3-0-A9-F. A
preferred embodiment also pertain toward the composition of SEQ ID NO:7, SEQ ID NO:8 and clone
1000849866_181-44-3-0-A9-F. Also preferred are polypeptides fragments having a biological activity
10 as described herein and the polynucleotides encoding the fragments.

[0312] The protein of SEQ ID NO:8 is a splice variant of the sequence of Alpha 1
Microglobulin (swissprot accession number P02760). The 221 amino acid lipoglobulin protein
displays a lipocalin motif and belongs to the lipocalin superfamily. These extra-cellular proteins bind
and transport small hydrophobic ligands such as steroids, bilins, retinoids, and lipids. Lipocalins
15 function in a variety of processes including nutrient transport, cell growth regulation, immune response
and prostaglandin synthesis.

[0313] Lipoglobulin is a plasma glycoprotein. The protein carries a set of chromophores
which gives the protein a yellow-brown color and an extremely heterogeneous charge. Lipoglobulin
can associate covalently with many plasma proteins like factor IX Zutphen, Factor XII Tenri and
20 several protein C mutants. Lipoglobulin is translated in the liver cell and secreted to the blood.
Lipoglobulin is found in many physiological fluids including plasma, urine, and cerebrospinal fluid.
The protein appears not only as a free monomer but also in complexes with IgA and albumin.
Lipoglobulin is also involved in anti-inflammatory and immunosuppressive activities. Lipoglobulin is
one of the positive acute-phase proteins, in that circulating levels of lipoglobulin increase in response
25 to stress and inflammatory stimulation. Lipoglobulin accumulates at sites of inflammation where it
inhibits platelet and neutrophil activation and inhibits phagocytosis. The immunomodulatory
properties of lipoglobulin are due to glycosylation. Lipoglobulin is 40 % carbohydrate, making it
unusually acidic and soluble. The glycosylation pattern of lipoglobulin changes during acute-phase
response, and deglycosylated lipoglobulin has no immunosuppressive activity. Lipocalins are used as
30 diagnostic and prognostic markers in a variety of disease states. The plasma level of Lipoglobulin can
be monitored during pregnancy and in diagnosis and prognosis of conditions including cancer
chemotherapy, renal dysfunction, myocardial infarction, arthritis, and multiple sclerosis.

[0314] An additional preferred embodiment of the invention is a method of binding lipoglobulin polypeptides with a lipoglobulin-specific antibody or lipoglobulin-binding fragment thereof. This method comprises the step of contacting a lipoglobulin polypeptide with a lipoglobulin-binding antibody or lipoglobulin-binding fragment thereof under conditions that allow binding. This
5 method may be applied to detection and purification of lipoglobulin. These aspects are discussed in detail herein.

[0315] An embodiment of the present invention relates to compositions comprising Lipoglobulin polypeptides. The method of producing Lipoglobulin polypeptides comprises the steps of: i) transfecting a host cell with a recombinant expression vector comprising a polynucleotide of the
10 present invention, and ii) purifying the produced protein. The purification of the protein can be done following any technique well known to those skilled in the art. Preferably, an antibody directed against Lipoglobulin or part thereof, preferably, an antibody directed against the C-terminal sequence of Lipoglobulin polypeptide, may be bound to a chromatographic support to form an affinity chromatography column.

15 [0316] In another embodiment, the invention provides methods and compositions for detecting the level of Lipoglobulin mRNA expression. Quantification of mRNA levels of Lipoglobulin may be useful for the diagnosis or prognosis of diseases associated with an altered expression of the protein of the invention. Conditions, diseases or disorders associated with altered expression include, but are not limited to, certain cancers, cystic fibrosis, ulcer, clotting disorders such
20 as hemophilia, inflammation and immune based disorders, acquired immunodeficiency syndrome (AIDS), autoimmune disorders such as arthritis, fertility disorders, and hypothyroidism.

[0317] Assays for the detection and quantification of the mRNA of the protein of the invention are well known in the art (see, for example, Maniatis, Fitsch and Sambrook, Molecular Cloning; A Laboratory Manual (1982), or Current Protocols in Molecular Biology, Ausubel, F.M. et
25 al. (Eds), Wiley & Sons, Inc.). For example, the nucleic acid molecule or probe may be labeled by standard methods and added to a biological sample from a patient under conditions of formation of hybridization complexes. After an incubation period, the sample is washed and the amount of label (or signal) associated with hybridization is quantified and compared with a standard value. If the amount of label in the patient sample is significantly altered in comparison to the standard value, then the
30 presence of the associated condition, disease or disorder is indicated. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies and in clinical trial and to monitor the treatment of an individual patient. Once the presence of a condition is established and a treatment protocol is initiated, diagnostics assays may be repeated on a regular basis to

determine if the level of expression in the patient begins to approximate that which is observed in a normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

[0318] In another embodiment, the Lipoglobulin polypeptide or fragment thereof may be used
5 to diagnose, disorders characterized by the presence of substrates, for example, Factor IX Zutphen,
Factor XII Tenri and several protein C mutants. Such disorders include but are not limited to, factor
XII deficiency and haemophilia B. [0319] In such a method, the Lipoglobulin polypeptide or
fragment thereof is used in assays and diagnostic kits for the identification and/or quantification of
substrates such as Factor IX Zutphen, Factor XII Tenri, and protein C mutants in a biological sample.

10 [0320] In another embodiment, the invention relates to methods and compositions for
detecting and quantifying the level of the protein of the invention present in a particular biological
sample. These methods are useful for the diagnosis or prognosis of diseases associated with altered
levels of the protein of the invention. Diagnostic assays to detect the protein of the invention may
comprise a biopsy, in situ assay of cells from organ or tissue sections, or an aspirate of cells from a
15 tumor or normal tissue. In addition, assays may be conducted upon cellular extracts from organs,
tissues, cells, urine, or serum or blood or any other body fluid or extract.

[0321] Detection and quantification of Lipoglobulin polypeptides using either specific
polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include
enzyme-linked immunoabsorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence
20 activated cell sorting (FACS). Example of an antibody is described in U.S. Patent 6,153,192, which
disclosure is incorporated by reference in its entirety.

[0322] In one embodiment, Lipoglobulin or a preferred fragment may be administered to a
subject to treat or prevent a condition associated with altered expression or activity of the protein of
the invention. Examples of such conditions include, but are not limited to, those described above.

25 [0323] In another embodiment, a pharmaceutical composition comprising a substantially
purified Lipoglobulin polypeptide in conjunction with a pharmaceutical carrier may be administered to
a subject to treat or prevent a condition associated with altered expression or activity of the
endogenous protein including, but not limited to, those provided above.

[0324] In another embodiment, a substance which modulates the activity of Lipoglobulin may
30 be administered to a subject to treat or prevent a condition associated with altered half life, expression,
or activity of Lipoglobulin polypeptides including, but not limited to, those listed above. In one
aspect, an antibody which specifically binds Lipoglobulin may be used as a targeting and delivery

mechanism for bringing a pharmaceutical agent to cells that express Lipoglobulin (i.e., hepatocytes). In an additional embodiment, a vector capable of expressing Lipoglobulin or a preferred fragment may be administered to a subject to treat or prevent a condition associated with altered half life, expression, or activity of the protein of the invention including but not limited to, those listed above.

5 [0325] In a still further embodiment, a vector of expressing the polynucleotide encoding Lipoglobulin or a fragment thereof may be administered to a subject to treat or prevent a condition associated with altered half life, expression, or activity of the protein of the invention including but not limited to, those described above.

[0326] In another embodiment, the invention relates to a method of screening for compounds 10 that bind a Lipoglobulin polypeptide or fragment thereof. Such method comprises the step of labeling the ligand molecule to be tested with a detectable label, such as a fluorescent, radioactive, or enzymatic tag, placing the ligand molecule to be tested in contact with Lipoglobulin polypeptide, or a fragment thereof under conditions which permit specific binding to occur, removing of non-specifically bound molecules, and detecting bound molecules using appropriate means. In a preferred 15 embodiment, the proteins of the invention or part thereof may be used to identify and/or quantify substrates using any techniques known to those skilled in the art. To find substrates, the proteins of the invention, or part thereof, or derivative thereof, may be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The 20 formation of binding complexes, between the proteins of the invention, or part thereof, or derivative thereof, and the agent being tested, may be measured by methods well known to those skilled in the art such as the BIACore (Upsala, Sweden). Antagonists or inhibitors of the proteins of the invention may be produced using methods which are generally known in the art, including the screening of libraries of pharmaceutical agents to identify those which specifically bind the protein of the invention. 25 Another technique for drug screening which may be used provides for high throughput screening of compounds having suitable binding affinity to the protein of the invention. For example, antagonists of Lipoglobulin polypeptide will prevent the inhibition of platelets and neutrophils by Lipoglobulin polypeptides.

Protein of SEQ ID NO:10 (Internal designation Clone 164341_117-001-5-0-E2-F)

30 [0327] The cDNA of clone 164341_117-001-5-0-E2-F (SEQ ID NO:9) encodes peroxisomal beta ketothiolase (PBK) of SEQ ID NO:10, comprising the amino acid sequence:

[0328] MQRLQVVLGHLRGPADSGWMPQAAPCLSGAPQASAADV VVVHGRRTAICR AGRGGFKDTPDELLSAVMTAVLKDVNLRPEQLGDICVG NVLQPGAGAIMARIAQFLSDIPET VPLSTVNRQCSSGLQAVASIAGWSPCPWLTEGTLEILLRA. Accordingly, it will be appreciated that all characteristics and uses of the polypeptides of SEQ ID NO:10 described throughout the present application also pertain to the polypeptides encoded by the nucleic acids included in Clone 164341_117-001-5-0-E2-F. In addition, it will be appreciated that all characteristics and uses of the polynucleotides of SEQ ID NO:9 described throughout the present application also pertain to the nucleic acids comprising the human cDNA in Clone 164341_117-001-5-0-E2-F. A preferred embodiment of the invention is directed toward the compositions of SEQ ID NO:9, SEQ ID NO:10, and clone 164341_117-001-5-0-E2-F. Also preferred are polypeptide fragments having a biological activity as described herein and the polynucleotides encoding the fragments.

[0329] The protein of SEQ ID NO:10, PBK, is a splice variant of the sequence of 3-Ketoacyl-CoA Thiolase protein (swissprot accession number P09110). The 153 amino acid protein of PBK displays one membrane-spanning segment: CSSGLQAVASIAGWSPCPWLT. Accordingly, some embodiments of the present invention relate to polypeptides comprising the transmembrane domain. Finally, the protein of the invention displays a thiolase domain spanning the sequence: APQASAADV VVVHGRRTAICRAGRGFKDTPDELLSAVMTAVLKDVNLRPEQLGDICVG NVLQPGAGAIMARIAQFLSDIPETVPLSTVNRQCSSGLQAVASIAGWSPCPWLTEGTLEIL Accordingly, a preferred embodiment of the present invention comprises the amino acids of the thiolase domain and polynucleotides encoding the same.

[0330] Living organisms are exposed to a number of different fatty acids and their various derivatives arising either via endogenous synthesis or from exogenous sources. These hydrophobic compounds can have specific metabolic, structural or endocrine functions before elimination, which can involve metabolism to CO₂ or to more polar lipid metabolites allowing their excretion. Quantitatively, one of the major pathways metabolizing fatty acids is beta-oxidation and PKB is involved in this pathway.

[0331] PBK is a ketoacyl-CoA thiolase involved in the formation of an acyl-enzyme intermediate and as such, plays a role in fatty acid metabolism, cellular vesicle transport and maintenance of the cytoarchitecture, cellular proteolysis, endocytosis, signal transduction, lysosomal storage, cell proliferation and differentiation, immune and inflammatory response. The enzyme's substrates are compounds preferably containing an ester bond, preferably a thiol ester bond, more preferably an acyl thioester bond.

[0332] PBK mediates lipid homeostasis in target organs like liver, adipose tissue and is regulated by peroxisome proliferator activated receptors (PPAR alpha and PPAR gamma). Deficiency of PBK leads to very low peroxisomal beta-oxidation activity and the accumulation of very-long-chain fatty acids and intermediates in the biosynthesis of bile acids. Such peroxisomal disorders are 5 peroxisome-deficient Zellweger syndrome and Rhizomelic Chondrodysplasia Punctata.

[0333] An additional preferred embodiment of the invention is a method of binding PBK polypeptides with a PBK-specific antibody or PBK-binding fragment thereof. This method comprises the step of contacting a PBK polypeptide with a PBK-binding antibody or PBK-binding fragment thereof under conditions that allow binding. This method may be applied to detection and purification 10 of PBK, as well as modification of PBK function. These aspects are discussed in detail herein.

[0334] A preferred aspect of the invention is a host cell recombinant for polynucleotides encoding a PBK polypeptide or a biologically active fragment thereof. An additional preferred aspect is a host cell recombinant for polynucleotides capable of directing PBK expression. Preferably, the polynucleotides capable of directing PBK expression are located in the 5' regulatory region of the 15 PBK gene. Further preferably, these polynucleotides are located within 500 base pairs of the PBK coding region. These polynucleotides preferably comprise a promoter sequence. Techniques known in the art for introducing polynucleotide sequences to endogenous sequences are described in U.S. Patent 5641670 and PCT WO9629411, which disclosures are hereby incorporated by reference in their entireties. PBK protein produced by said host cell may be used for in vitro detection and purification 20 methods as well as diagnosis and in vivo applications.

[0335] In one embodiment, the hydrolytic activity of the protein of the invention or part thereof may be assayed using any of the assays known to those skilled in the art including those described in US patents 5,445,942, which disclosure is hereby incorporated in its entirety.

[0336] Another embodiment of the invention relates to compositions and methods using the 25 protein of the invention or part thereof to label peroxisomes, in order to visualize any change in number, topology or morphology of this organelle, for example in association with beta-oxidation peroxisomal disorders such as, but not limited to Zellweger syndrome, Pseudo Zellweger, Rhizomelic Chondrodysplasia Punctata, X-linked adrenoleukodystrophy (ALD), neonatal adrenoleukodystrophy, pseudoneonatal adrenoleukodystrophy, acyl-coA deficiency, bifunctional enzyme deficiency, 30 Refsum's disease, DHAP acyl transferase deficiency, hyperpipecolatemia and acatalasemia.

[0337] For example, the protein may be rendered easily detectable by inserting the cDNA encoding the protein of the invention into a eukaryotic expression vector in frame with a sequence

encoding a tag sequence. Eukaryotic cells expressing the tagged protein of the invention may also be used for the *in vitro* screening of drugs or genes capable of treating any beta-oxidation peroxisomal disorders.

[0338] An embodiment of the invention provides for a method of screening test substances for modulators of PBK expression. This method comprises the steps of: i) contacting a cell with a test substance; and ii) comparing PBK expression in the cell after exposure to the test substance to that of an unexposed control cell. PBK expression is determined by methods common to the art or included herein, by detecting PBK polynucleotides or polypeptides. An example of this method comprises the steps of: i) culturing two equivalent cell samples; ii) adding a test substance to one of the cultures and not the other; iii) harvesting both cultures at a specified time; iv) purifying the mRNA from each sample of cells; v) comparing the level of PBK mRNA in each sample by Northern blot, RTPCR, or another method common to the art. The invention provides for design and use of specific polynucleotide probes and primers, as discussed herein. An additional example comprises the steps of: i) having two equivalent cultures of cells; ii) adding a test substance to one of the cultures and not the other; iii) harvesting both cultures; iv) purifying the protein from each sample of cells; v) comparing the level of PBK polypeptides in each sample by Western blot, immunohistochemistry, or another method common to the art. The invention provides for design and use of specific antibodies and antibody fragments, as discussed herein.

[0339] Agents which modulate the expression or activity of the PBK of the subject invention include, but are not limited to antisense oligonucleotides, ribozymes, drugs, and antibodies. These agents may be made and used according to methods well known in the art. Also, agents which increase PBK expression are hypolipidemic compounds. Moreover, the protein of the invention, or biologically active fragments thereof, may be used in screening assays for therapeutic compounds. A variety of drug screening techniques may be employed. In this aspect of the invention, the protein or biologically active fragment thereof, may be free in solution, affixed to a solid support, recombinantly expressed on, or chemically attached to, a cell surface, or located intracellularly. The formation of binding complexes, between the protein of the invention, or biologically active fragments thereof, and the compound being tested, may be measured by methods well known to those skilled in the art, like, but not limited to, the BIACore (Upsala, Sweden). Another technique for drug screening which may be used provides for high throughput screening of compounds having suitable binding affinity to the protein of the invention as described in published PCT application WO84/03564, and incorporated herein by reference in its entirety.

[0340] Another embodiment of the subject invention provides compositions and methods of selectively increasing the activity of the protein of the invention. Activation of PBK allows for the successful treatment and/or management of diseases or biochemical abnormalities associated with PBK activity. Agonists, able to increase the expression or the activity of the protein of the invention, 5 are useful in the treatment of diseases associated with decreased peroximal beta-oxidation of fatty acids, for example, Pseudo Zellweger syndrome and Rhizomelic Chondrodysplasia Punctata (RCDP). Preferred agonists include hypolipidemic compounds. Each hypolipidemic compound may additionally be excluded individually or as a group.

[0341] In another embodiment, the present invention includes the use of PBK polypeptides, or 10 fragments having a desired biological activity to treat or ameliorate a condition in an individual. For example, the condition may be Pseudo-Zellweger syndrome or Rhizomelic Chondrodysplasia Punctata (RCDP). In such embodiments, a PBK polypeptide or a fragment thereof, is administered to an individual in whom it is desired to increase any of the activities of PBK. A PBK polypeptide or fragment thereof may be administered directly to the individual or, alternatively, a nucleic acid 15 encoding the a PBK polypeptide or a fragment thereof may be administered to the individual. For example, a polynucleotide encoding a PBK polypeptide or a fragment is comprised in a recombinant expression vector which can be administered to a patient suffering from Pseudo-Zellweger syndrome or Rhizomelic Chondrodysplasia Punctata (RCDP). Preferred expression vectors include viral vectors, especially adenoviral and lentiviral vectors. Alternatively, an agent which increases the activity of 20 PBK polypeptides may be administered to the individual. Preferred agonists are hypolipidemic compounds. Specific hypolipidemic compounds may additionally be excluded individually or as a group.

[0342] In such a method, an agent that increases PBK activity can be processed in accordance with conventional methods of pharmacy to produce medicinal agents for administration to patients. 25 Thus, the pharmaceutical composition comprising an agent that increases PBK activity may be made up in a solid form (e.g. granules for oral administration, powders for inhalation) or in a liquid form (e.g. solutions for oral administration or for injection). Effective doses of the compositions of the present invention for treating a patient suffering from disorders like, but not limited to, Pseudo-Zellweger syndrome or Rhizomelic Chondrodysplasia Punctata (RCDP), can be determined according 30 to the relevant techniques. Additional agents may be identified by contacting a PBK polypeptide or a cell or preparation containing a PBK polypeptide with a test agent and assaying whether the test agent increases the activity of the protein. For example, the test agent may be a chemical compound or a polypeptide or peptide.

[0343] In one embodiment, the subject method utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the PBK polypeptide or biologically active fragments thereof. The transformed cells may be viable or fixed. Drugs or compounds which are candidates for the activation of the protein of the invention, or biologically active fragments thereof, are screened against such transformed cells in binding assays well known to those skilled in the art. Alternatively, assays such as those taught in Geysen H. N., WO Application 84/03564, published on Sep. 13, 1984, and incorporated herein by reference in its entirety, may be used to screen for peptide compounds which demonstrate binding affinity for, or the ability to activate, PBK polypeptides or biologically active fragments thereof.

[0344] Another embodiment of the invention relates to composition and methods using polynucleotide sequences encoding the protein of the invention or part thereof to establish transgenic model animals (*D. melanogaster*, *M. musculus*), by any method familiar to those skilled in the art. By modulating *in vivo* the expression of the transgene with drugs or modifier genes (activator or suppressor genes), animal models can be developed that mimic human peroxisome-associated disorders such as Zellweger syndrome, Rhizomelic Chondrodysplasia Punctata, X-linked adrenoleukodystrophy (ALD), acyl-coA deficiency, bifunctional enzyme deficiency, Refsum's disease, DHAP acyl transferase deficiency and acatalasemia. These animal models would thus allow the identification of potential therapeutic agents for treatment of the disorders. In addition, recombinant cell lines derived from these transgenic animals may be used for similar approaches *ex vivo*.

[0345] In another embodiment, the invention relates to methods and compositions for detecting and quantifying the level of the protein of the invention present in a particular biological sample. These methods are useful for the diagnosis or prognosis of diseases associated with altered levels of the protein of the invention like, but not limited to, pseudo-Zellweger syndrome. Diagnostic assays to detect the protein of the invention may comprise a biopsy, *in situ* assay of cells from organ or tissue sections, or an aspirate of cells from a tumor or normal tissue. In addition, assays may be conducted upon cellular extracts from organs, tissues, cells, urine, or serum or blood or any other body fluid or extract.

[0346] Assays for the quantification of the protein PBK may be performed according to methods well known in the art. Typically, these assays comprise contacting the sample with a ligand of the protein of the invention or an antibody (polyclonal or monoclonal) which recognizes the protein of the invention or a fragment thereof, and detecting the complex formed between the protein of the invention present in the sample and the ligand or antibody. Fragments of the ligands and antibodies

may also be used in the binding assays, provided these fragments are capable of specifically interacting with the protein of the subject invention. Further, the ligands and antibodies which bind to the protein of the invention may be labeled according to methods known in the art. Labels which are useful in the subject invention include, but are not limited to, enzymes labels, radioisotopic labels, 5 paramagnetic labels, and chemiluminescent labels. Typical techniques are described by Kennedy, J. H., *et al.* (1976) *Clin. Chim. Acta* 70:1-31, which disclosure is hereby incorporated by reference in its entirety; and Schurs, A. H. *et al.* (1977) *Clin. Chim. Acta* 81: 1-40, which disclosure is hereby incorporated by reference in its entirety). For example, PBK will be detected in fibroblasts of control individuals but not of Zellweger patients by immunoblot.

10 [0347] In another embodiment, an array of oligonucleotides probes comprising the nucleotides encoding PBK polypeptides or fragments thereof can be constructed to conduct efficient screening of e.g., genetic mutations. A microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a 15 disorder, to diagnose a disorder, and to develop and monitor the activities of therapeutic agents (see for example: Chee, M. *et al.*, *Science*, 274:610-614 (1996), which disclosure is hereby incorporated by reference in its entirety). Genetic variants, mutations, and polymorphisms of PBK gene are related to peroxisomal beta-oxidation disorders such as Rhizomelic Chondrodysplasia Punctata.

Protein of SEQ ID NO:12 (Internal designation Clone 1000837037_228-43-2-0-C3-F)

20 [0348] The cDNA of Clone 1000837037_228-43-2-0-C3-F (SEQ ID NO: 11) encodes myeloidin of SEQ ID NO:12, comprising the amino acid sequence: MPFSHLSTYSLVVVMAAVVLCTAQVQVVTQDEREQLYTTASLKCSLQNAQEALIVTWQKK KAVSPENMVTSENHGVIQPAYDKINITQLGLQNSTITFWNITLEDEGCYMCLFNTFGFGKI SGTACLTYYVQPIVSLHYKFSEDLHNITCSATARPAPMVFWKVPNSGIENSTVTLHPNGTTSV 25 TSILHIKDPKNQVGKEVICQVLHLGTVTDFKQTVNKGYWFSVPLLLSIVSLVILLVLISILLYWK RHRNQDREP. Accordingly, it will be appreciated that all characteristics and uses of the polypeptides of SEQ ID NO:12 described throughout the present application also pertain to the polypeptides encoded by the nucleic acids comprising the human cDNA in Clone 1000837037_228-43-2-0-C3-F . In addition, it will be appreciated that all characteristics and uses of the polynucleotides of SEQ ID 30 NO:11 described throughout the present application also pertain to the nucleic acids included in Clone 1000837037_228-43-2-0-C3-F. A preferred embodiment of the invention is directed toward the compositions of SEQ ID NO:12, SEQ ID NO:11, and Clone 1000837037_228-43-2-0-C3-F. Preferred myeloidin polypeptides for uses in the methods described below include the polypeptides comprising

the amino sequence of:
QVQVVTQDEREQLYTTASLKCSLQNAQEALIVTWQKKAVSPENMVTFSENHGVVIQPAYK
DKINITQLGLQNSTITFWNITLEDEGCYMCFLNTFGFGKISGTACLTIVYVQPIVSLHYKFSEDHL
NITCSATARPAVMFWKVPRSGIENSTVTLHPNGTTSVTSILHIKDPKNQVGKEVICQVLHLG
5 TTVDFKQTVNKGYWFSVPILLSIVSLVILLVLISILLYWKRHRNQDREP. Also preferred are
polypeptide fragments having a biological activity as described herein and the polynucleotides
encoding the fragments.

[0349] The protein of the present invention, myeloidin, is a novel splice variant of the brain OX-2 protein (Genbank accession number P41217). Myeloidin is 262 amino acids long. OX-2 is 274 10 amino-acids long and differs from myeloidin at both the amino-terminal and the carboxyl-terminal extremities. Another known splice variant of OX-2, my033, is 269 amino acids long and differs from myeloidin at the amino-terminal extremity. Myeloidin displays a signal peptide (MPFSHLSTYSLVVVMAAVVLCTA), one transmembrane domain (VPLLLSIVSLVILLVLISILLYW) and two immunoglobulin (Ig) domains 15 (TASLKCSLQNAQEALIVTWQKKAVSPENMVTFSENHGVVIQPAYKDKINITQLGLQNSTITF
WNITLEDEGCYMCFL and EDHLNITCSATARPAVMFWKVPRSGIENSTVTLHPNGTTSVTSILHIKDPKNQVGKEVICQV
). Myeloidin is expressed by a wide variety of cells, including those of the central nervous system (CNS).

20 [0350] Myeloidin is a cell surface protein that delivers an inhibitory signal for myeloid lineage cells. Myeloidin binds to a receptor specifically located on myeloid lineage cells. In peculiar, myeloidin delivers an inhibitory signal to brain microglial cells. Microglial cells regulate the regenerative state and remodeling of the brain by producing a variety of cytotoxic and neurotrophic molecules when activated. More generally, myeloidin is involved in immune suppression by 25 inhibiting cytotoxic cells.

[0351] An embodiment of the invention is directed to a composition comprising a myeloidin polypeptide sequence of SEQ ID NO:12.

[0352] A further embodiment of the invention is directed to a composition comprising a myeloidin polypeptide fragment having biological activity of delivering an inhibitory signal for 30 myeloid lineage cells.

[0353] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence of SEQ ID NO:11 encoding a myeloidin polypeptide.

[0354] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence encoding a myeloidin polypeptide fragment having biological activity of delivering an inhibitory signal for myeloid lineage cells.

[0355] A further embodiment of the invention is directed to a composition comprising an antibody recognizing a myeloidin polypeptide sequence of SEQ ID NO:12 or a myeloidin polypeptide fragment having same biological activity. Preferably, the antibody binds to myeloidin but not to OX-2 or to my033. As further used herein, an *anti-myeloidin antibody* refers to an antibody that specifically binds either to a myeloidin polypeptide or to a myeloidin polypeptide fragment having the same biological activity. An anti-myeloidin antibodies may be an entire molecule or an antigen binding fragment thereof.

[0356] The present invention also relates to a method of binding an anti-myeloidin antibody to a myeloidin polypeptide comprising the step of: contacting a myeloidin polypeptide with said antibody under conditions that allow binding to take place. Such conditions are well known to those skilled in the art. Additionally, the present invention also relates to a method of binding an anti-myeloidin antibody to a cell expressing a myeloidin polypeptide or fragment thereof comprising the step of: contacting a cell expressing a myeloidin polypeptide or part thereof with said antibody under conditions that allow binding to take place. Such methods of binding an anti-myeloidin antibody to a myeloidin polypeptide are for example useful for purifying myeloidin polypeptides.

[0357] Another embodiment relates to a method of producing myeloidin polypeptides comprising the steps of: i) transfecting a host cell with a recombinant expression vector comprising a polynucleotide of the present invention, and optionally ii) purifying the produced protein. The purification of the protein can be done following any technique well-known to those skilled in the art. Preferably, an anti-myeloidin may be bound to a chromatographic support to form an affinity chromatography column. Alternatively, myeloidin polypeptides may be produced by a method comprising the step of: i) transfecting a host cell with a polynucleotide capable of directing myeloidin expression. Preferably, the polynucleotides capable of directing myeloidin expression are located in the 5' regulatory region of the myeloidin gene. Further preferably, these polynucleotides are located within 500 base pairs of the myeloidin coding region. These polynucleotides preferably comprise a promoter sequence. Techniques known in the art for introducing polynucleotide sequences to endogenous sequences are described in U.S. Patent 5641670 and PCT WO9629411, which disclosures are hereby incorporated by reference in their entireties.

[0358] An additional aspect of the previous embodiment relates to methods of producing myeloidin *in vivo* in a mammal. Preferred such method is a method of genetically modifying donor cells by gene transfer of myeloidin polynucleotides for grafting the cells into the central nervous as described in U.S. Patent 5,762,926, which disclosure is incorporated by reference in its entirety.

5 Preferably, donor cells are neurons or fibroblasts. Most preferably, such methods are applied to a mammal suffering from neurodegenerative disorders. Preferred mammals are mice and humans.

[0359] One embodiment of the present invention is directed to methods of isolating or purifying microglial cells using myeloidin polypeptides. The study and use of isolated microglial cells and purified microglial cell populations (including, for example, characterizing the interaction these 10 cells *in vitro* or *in vivo* with drugs and drug candidates) provides information useful in treating neurological disorders. One method for obtaining a cell population enriched in microglial cells comprises the steps of: i) contacting a preparation comprising microglial cells with a myeloidin polypeptide, wherein myeloidin is immobilized on a matrix prior to or after contacting; and ii) removing non-adherent cells, thereby producing cell population enriched in microglial cells. For 15 example, the matrix may be plastic and the myeloidin polypeptide may be immobilized by adsorption. Another method for isolating or purifying microglial cell populations using compositions comprising using an myeloidin polypeptide comprises the steps of: i) labeling by standard methods the myeloidin polypeptide with a molecule that can be used to provide a detectable signal, ii) contacting the labeled myeloidin polypeptide with a cell under conditions that allow protein binding, and iii) sorting the cell 20 based on the presence or absence of the detectable signal. Sorting and detection are preferably accomplished by a sorting apparatus, e.g. a fluorescence-activated cell sorter or a magnetic activated cell-sorting apparatus. Preferably, the preparations comprising microglial cells are prepared from brain (e.g., from rat or mouse brain) and may be from a normal mammal or from a mammal with a neurological disorder. Preparation of cell suspension tissue containing microglia can be performed 25 according to any method well-known to those skilled in the art, e.g., the methods described in PCT application WO 01/51618, which disclosure is hereby incorporated by reference in its entirety.

[0360] Another preferred embodiment relates to methods of screening for substances that increase or decrease myeloidin expression comprising the steps of: i) contacting a cell with a test substance; and ii) comparing myeloidin expression in the cell after exposure to the test substance to 30 that of an unexposed control cell. Preferably, the test substance modifies the expression of myeloidin in a specific cell type while not in others. Most preferably, the test substance modifies myeloidin expression specifically in neural cells.

[0361] A further embodiment of the present invention is directed to substances that increase or decrease myeloidin activity and to a method of screening for such substances comprising the steps of: i) contacting a cell with a test substance, ii) determining myeloidin activity, iii) comparing myeloidin activity in the cell after exposure to that of an unexposed control cell, and iv) determining 5 the ratio of myeloidin activity in an exposed cell relative to the activity in an unexposed cell. Preferably, myeloidin activity is studied in a neural cell. Myeloidin activity can be determined by studying microglial-mediated activation of astrocytes, *i.e.*, to determine its ability to prevent a reactive astrocyte morphology in culture. Myeloidin activity can also be determined by measuring cytokine production as described in PCT application WO 99/24565, which disclosure is hereby incorporated by 10 reference in its entirety.

[0362] As further used herein, substances that increase myeloidin expression or myeloidin activity are defined as myeloidin agonists. As further used herein, substances that decrease myeloidin expression or myeloidin activity are defined as myeloidin antagonists. The term antagonist comprises anti-myeloidin antibodies that inhibit the biological activity of myeloidin polypeptides.

[0363] An embodiment of the present invention relates to methods of using myeloidin antagonists of the present invention to induce or to enhance activation of myeloid lineage cells. Such methods comprise the step of: contacting a cell, a tissue, or an individual with a composition comprising a myeloidin antagonist and optionally a pharmaceutical carrier. In a preferred embodiment, such methods for activating myeloid lineage cells are directed toward brain microglial 15 cells. Such methods can be used to destroy invading microorganisms, to promote neuroregeneration, or to remove potentially deleterious debris and promote tissue repair consecutively to brain injury. Infections that can be treated or reduced in severity by activating microglial cells include but are not limited to brain abscess due to, e.g., *Aspergillus* CNS infections, meningitis due to, e.g., *Cryptococcus* CNS infections, neuropsychiatric disorders associated with streptococcal infection (PANDAS), and 20 influenza virus-associated encephalopathy.

[0364] Another embodiment of the present invention relates to methods of using the polypeptides and the polynucleotides of the present invention to prevent or to reduce activation of myeloid lineage cells. Such methods comprise the step of: contacting a cell, a tissue, or an individual with a composition comprising a myeloidin polypeptide or a myeloidin agonist, and optionally a 25 pharmaceutical carrier. In a preferred embodiment, such methods for preventing or reducing myeloid lineage cell activation are directed toward brain microglial cells. Such methods can be used to treat or to reduce in severity neural inflammation, and more specifically neurodegenerative disorders and pathological inflammations consecutive to brain injury. Neurodegenerative disorders associated with

neural inflammation include but are not limited to stroke, Parkinson's disease, Alzheimer's disease, cerebellum-type Creutzfeldt-Jakob disease, and Huntington's disease.

[0365] In another preferred embodiment, such methods for preventing or reducing myeloid lineage cell activation are directed for preventing or treating inflammation associated with 5 autoimmune diseases. Autoimmune diseases that may be treated or prevented according to the present invention include, but are not limited to, multiple sclerosis, type 1 insulin-dependant diabetes mellitus, and rheumatoid arthritis.

[0366] Effectiveness of compositions and methods of the present invention to modulate microglial activity can be verified *in vitro* by studying the effects of the compositions of the present 10 invention on the morphology of astrocytes. Effectiveness of compositions and methods of the present invention to treat a given disease associated with microglial activation can be verified using models that depend upon the disease that is treated. For example, one model described in US Patent 6,191,154, which disclosure is hereby incorporated by reference in its entirety, can be used as a tissue model for the sequence of events following trauma in the nervous system.

15 [0367] The quantity of myeloidin polypeptides, anti-myeloidin antibodies, myeloidin agonists or myeloidin antagonists that is administered for treating a patient suffering from diseases described above can be determined according to the particular application and the potency of the active component. For preparing pharmaceutical compositions from the compounds of the present invention, pharmaceutically acceptable carriers can be in any suitable form (e.g., solids, liquids, gels). The 20 present invention contemplates a variety of techniques for administration of the therapeutic compositions. Suitable routes include, but are not limited to, oral, rectal, or transdermal administration, and intradermal, subcutaneous or intravenous injections. Indeed, it is not intended that the present invention is limited to any particular administration route. The assessment of the clinical features and the design of an appropriate therapeutic regimen for the individual patient are ultimately 25 the responsibility of the prescribing physician. Furthermore, the compositions of the invention may be administered alone or in combination with other known agents treating the disorders treated above.

Protein of SEQ ID NO:14 (Internal designation Clone 101005_105-020-4-0-H11-F)

[0368] The cDNA of Clone 101005_105-020-4-0-H11-F (SEQ ID NO:13) encodes protein vACT of SEQ ID NO:14, comprising the amino acid sequence:

30 MERMLPLTLGLLAAGFCPAVLCHPNSPLDEENLTQENQDRGTHVDLGLASANVDFALSLYK
QLVLKAPDKNVIFSPLSISTALAFSLGAHNTTLTEILKGLKFNLTETSEAEIHQSFQHLLRTLN
QSSDELQLSMGNAMFVKEQLSLLDRFTEDAKRLYGSEAFATDFQDSAACKLINDYVKNGT

RGKITDLIKDLDQTMMVLVNYIFFKAKWEMPFDQDTHQSRFYLSKKWVMVPMMSLHHL
TIPYFRDEELSCTVVELKYTGNASALFILPDQDKMEEVEAMLLPETLKRWRDSLEFREIGELYL
PKFSISRDYNLNDILLQLGIEEAFTSKADLSGITGARNLAQSQVVKAVLDVFEEGTEASAATA
VKITLLSALVETRTIVRFNRPFLMIIVPTDTQNIFFMSKVTPNKQA. Accordingly, it will be
5 appreciated that all characteristics and uses of polypeptides of SEQ ID NO:14 described throughout
the present application also pertain to the polypeptides encoded by the nucleic acids comprising the
human cDNA in Clone 101005_105-020-4-0-H11-F. In addition, it will be appreciated that all
characteristics and uses of the polynucleotides of SEQ ID NO:13 described throughout the present
application also pertain to the nucleic acids included in Clone 101005_105-020-4-0-H11-F. A
10 preferred embodiment of the invention is directed toward the compositions of SEQ ID NO:13, SEQ ID
NO:14, and Clone 101005_105-020-4-0-H11-F. Also preferred are polypeptide fragments having a
biological activity as described herein and the polynucleotides encoding the fragments.

[0369] The cDNA of SEQ ID NO:13 is a novel polymorphic variant of the human alpha 1
anti-chymotrypsin (ACT) protein named vACT, encoded by a gene located on chromosome 14,
15 specifically at position 14q31-q32.3. The cDNA of SEQ ID NO:13 encodes a 423 amino-acid protein
of SEQ ID NO:14.

[0370] Proteases are key components of a broad range of biological pathways and can be
classified into four groups according to their catalytic mechanisms: the serine, cysteine (thiol), aspartic
(carboxyl), and metalloproteases. Chymotrypsin is a member of a family of enzymes known as serine
20 proteases, so named because they have an unusually reactive serine residue at their active sites. vACT
belongs to the serpin serine protease inhibitor family. Serpins are irreversible suicide inhibitors of
proteases that have a central role in regulating proteolysis in diverse physiological processes such as
blood coagulation, fibrinolysis, complement activation, angiogenesis, apoptosis, inflammation,
neoplasia and viral pathogenesis. vACT neutralizes chymotrypsin by binding to its active site and
25 forming a stable complex. vACT is predominantly synthesized in the liver and specifically inactivates
neutrophil cathepsin G, mast cell chymase and pancreatic chymotrypsin. vACT is also an acute phase
protein since the plasma concentration of vACT increases within several hours during the
inflammatory response. Synthesis of vACT is tightly regulated by the net balance of neutrophil
cathepsin G and vACT at sites of inflammation/tissue injury. Alterations of a serpin which affect its
30 functional levels may result in pathology and cause specific clinical syndromes. For example,
individuals with vACT deficiency are susceptible to premature development of lung and liver diseases.
In addition, changes in the balance between serine proteases and vACT may lead to pathological states
similar to those associated with some neurodegenerative diseases such as Alzheimer's disease.

[0371] An embodiment of the invention is directed to a composition comprising a vACT polypeptide sequence of SEQ ID NO:14.

[0372] A further embodiment of the invention is directed to a composition comprising a vACT polypeptide fragment having biological activity.

5 [0373] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence of SEQ ID NO:13 encoding a vACT polypeptide.

[0374] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence encoding a vACT polypeptide fragment having biological activity.

[0375] In an embodiment, a vACT polypeptide or fragment thereof may be used to inhibit 10 contaminating proteases in a sample. This method comprises the step of adding a protease-inhibiting amount of vACT polypeptide to biological samples, under conditions that allow vACT activity to prevent degradation of protein samples. Preferred biological samples are cell cultures. Such a composition can be used alone or as a "cocktail" with other protease inhibitors. The advantage of 15 using a cocktail of protease inhibitors is that one is able to inhibit a wide range of proteases without knowing the specificity of any of the proteases. Using a cocktail of protease inhibitors also protects a protein sample from a wide range of proteases which may contaminate a protein sample in the future from a vast number of sources. For example, vACT polypeptides or fragments thereof are added to samples where proteolytic degradation by contaminating proteases is undesirable. Such protease inhibitor cocktails are widely used in assays to inhibit proteases susceptible of degrading a protein of 20 interest for which the assay is to be performed. Alternatively, vACT or part thereof may be bound to a chromatographic support, either alone or in combination with other protease inhibitor, using techniques well known in the art, to form an affinity chromatography column. A sample containing the undesirable protease is run through the column to remove the protease. Alternatively, the same methods may be used to identify new target proteases of vACT.

25 [0376] In another preferred embodiment, vACT polypeptides or fragments thereof may be used as an anti-microbial agent useful to inhibit exogenous proteases implicated in a number of infectious diseases including, but not limited to, bacterial and parasite-borne infections. For example, vACT inhibits growth of all strains of group A streptococci, including antibiotic-resistant strains. Accordingly, the present invention can be used to retard or inhibit the growth of certain microbes 30 either *in vitro* or *in vivo*. Such methods comprise the step of delivering an amount of vACT polypeptide able to inhibit serine proteases, including by administering purified protein or transfecting cells with a polynucleotide encoding the protein. *In vitro*, this method comprises the step of contacting

an effective amount of vACT polypeptide with a biological sample. Preferred biological samples are cell cultures. *In vivo*, this method comprises the step of contacting an effective amount of vACT polypeptide with a desired site in an individual. Preferred individuals are those at risk of microbial infection.

5 [0377] In a further embodiment, the present invention provides a method of producing a recombinant serpin capable of effectively modulating serine protease activity. Despite the availability of human alpha 1 anti-chymotrypsin and vACT from serum, quantities large enough for therapeutic uses have been unobtainable, due in large part to the limited availability of human serum. Consequently, there is a great need for other sources of alpha 1 anti-chymotrypsin and vACT to fill the
10 needs created by therapeutic uses. In one preferred embodiment, a milk animal can be used to produce vACT in the milk, thereby generating a significant amount of this particular protein after purification. The protein of the invention may be purified using any techniques known to those skilled in the art including those disclosed in the US Patent 6,268,487, which disclosure is hereby incorporated by reference in its entirety. Any type of animal that produces enough milk can be used in this aim such
15 as, but not limited to, sheep, goat, and cow. These animals can be generated with any method of targeting overexpression of vACT in the milk. Also in this embodiment, the protein of the invention can be produced in host cells that have been transfected with an appropriate expression vector comprising a nucleic acid sequence coding for vACT. The host cells are cultured under conditions whereby the nucleic acid sequence coding for this particular protein is expressed. After a suitable
20 amount of time for the product to accumulate, the protein is purified from the host cells or medium surrounding the cells. Introduction of an expression vector incorporating a nucleic acid sequence coding for vACT into a host cell can be performed in a variety of ways, such as but not limited to calcium or lithium chloride treatment, electroporation, and lipofection.

[0378] The present invention also provides animal models generated by modulating the
25 expression or activity of the present protein in one or more tissues of the animal. Such animals are useful for a number of purposes, because they represent an *in vivo* assay method for testing candidate molecules potentially useful for the treatment of various pathophysiological aspects of diseases specifically related to the activity of vACT. Study of the phenotype of such models can also allow the identification of additional human equivalent diseases caused by or linked with vACT deficiency.
30 These animals can be generated with any method of targeting overexpression or inactivation of vACT. Such models are extremely useful, e.g. in the assessment of candidate therapies and drugs for the treatment of inflammatory diseases and conditions.

[0379] A further embodiment of the present invention is to provide novel methods and compositions useful for the treatment of diseases and conditions exhibiting excessive activity of serine proteases and preferably chymotrypsin. vACT or part thereof may be used to inhibit proteases implicated in a number of diseases where cellular proteolysis occurs. Such diseases, characterized by 5 tissue degradation, include, but are not limited to, blood coagulation-related diseases, tumor invasion, infection, and inflammation. More preferably, the present invention is applied in the treatment of diseases associated with an excess level of cathepsin G, mast cells chymase or pancreatic chymotrypsin including but not limited to, chronic emphysema of the lungs, liver diseases, pancreatitis, cardiovascular diseases, and allergic reactions. The methods and compositions are also 10 useful for treatment of Alzheimer's and Parkinson's diseases. Such methods comprise the administration of a therapeutically-effective amount of vACT to mammals suffering from the disease or condition, where "effective amount" is meant a concentration of vACT capable of significantly decrease the activity of serine proteases. The compositions of the invention are preferably delivered to an individual in combination with a physiologically acceptable carrier, such as a saline solution or 15 other physiologically buffer suitable for administration to a patient. For treatment of skin inflammation, the compositions of the invention may be applied to the affected area in combination with a physiologically acceptable ointment or cream. The particular amount of the compositions of the invention that will be administered to the mammal for any particular condition will depend on the clinical condition of the patient, and other factors such as the weight, age, and route of delivery. Such 20 compositions can be administered by any suitable route including, but not limited to, intravenous, intramuscular, intraperitoneal, subcutaneous routes, and topically to an affected area of the skin or by absorption through epithelial or mucocutaneous linings such as nasal, oral, vaginal, rectal. For treatment purposes, the compositions may be administrated using any of the gene therapy methods known in the art to deliver the therapeutic bioactive molecules to targeted cells. These compositions 25 can comprise the protein of the invention, and, optionally, one or more other types of protease inhibitors, or any other compound of interest. This co-administration may be by simultaneous administration or by separate or sequential administrations. All of these additional components may be either obtained from natural sources or produced by recombinant genetic engineering techniques and/or chemical modification.

30 [0380] An additional preferred embodiment of the invention is a method of binding vACT polypeptides with a vACT-specific antibody or vACT-binding fragment thereof. This method comprises the step of contacting a vACT polypeptide with a vACT-binding antibody or vACT-binding fragment thereof under conditions that allow binding. This method may be applied to detection and

purification of vACT, as well as modifying vACT function. These aspects are discussed in detail herein.

[0381] In other embodiment, the current invention is used to diagnose diseases or disorders associated with altered expression or activity of the present protein. In particular, it is useful in diagnosing patients with abnormally low levels of vACT expression, which results in uncontrolled activity of target proteases. Examples of such diseases and disorders include, but are not limited to, lung and liver diseases, chronic obstructive pulmonary disease (COPD), inflammatory disorders associated with an excess level of cathepsin G, and Alzheimer's and Parkinson's diseases. The method includes the steps of: i) contacting a fluid or tissue sample with a compound capable of selectively binding vACT polypeptide or nucleic acids, and ii) detecting the level, or any other detectable property of vACT in the sample. Preferably, a difference in the level or other property in the sample relative to in a control sample indicates the presence of the disease or disorder, or of a propensity for developing the disease or disorder. Preferably, the fluid or tissue sample is obtained from an individual suspected of suffering from the disease or condition, or at risk of developing the disease or condition. For example, a polyclonal or monoclonal antibody or any immunologically active fragment thereof or a nucleic acid probe may be used. Preferably, the antibodies used in one such embodiment are specifically directed against the variant vACT polypeptides and do not recognize alpha 1 anti-chymotrypsin polypeptides. Detection can be carried out directly or indirectly with known immunohistological and immunofluorescing processes. For this, the reagents contained in the kit according to the invention can be directly labeled with generally known molecules, including, but not limited to, enzymes such as alkaline phosphatase and peroxidase and fluorescent dyes such as FITC, rhodamine, and Texas-Red. However, labeling can also occur indirectly by using secondary antibodies labeled with molecules such as biotin, digoxigenin or the like and are then detected with a secondary reagent. Also in this embodiment, diagnosis of such conditions may be facilitated by the identification of the variant vACT using well known PCR or RT-PCR techniques and in particular in with "real-time" PCR system. Alternatively, using such a method, the present invention provides a tool to correlate modulations in the expression of vACT with certain pathologies. Thus, the present invention provides a novel candidate gene for such conditions.

[0382] Since the regulation of serine proteases by their inhibitors are critical for the control of tissue destruction in the diseases described above, in a further embodiment, the present protein or part thereof provides an assay for the monitoring of markers *in vivo* for characterization of disease states. The invention thus includes test kits useful for the quantification of the amount of vACT in a biological sample. The kits comprise at least one immunological binding partner, e.g. a monoclonal or

polyclonal antibody specific for vACT and coupled to detectable markers. Preferably, the antibodies used in one such embodiment specifically bind vACT polypeptides and not alpha 1 anti-chymotrypsin polypeptides. In this embodiment, the application of such assays can be used to monitor the progress of therapy administered to treat these or other conditions. Further, the assays can be used as a measure 5 of toxicity, or during clinical testing of new drugs to assess the impact on tissue degradation. Thus the assays may be applied in any situation wherein the present invention can be used as an index of the condition, treatment, or effect of substances directly administered to the subject or to which the subject is exposed in the environment. Thus, the condition of a patient can be monitored continuously and the quantified amount of such proteins measured in the pathological sample can be compared with the 10 amount quantified in a biological sample of a normal individual or with the previous analysis of the same patient. In this embodiment, this marker can be measured effectively in plasma, serum or blood, by any suitable method, including immunoassays. It can also preferably be measured in tissues and fluids recovered from inflammatory sites.

Protein of SEQ ID NO:16 (Internal designation 500743419_188-281-3-0-H5-F)

15 [0383] The cDNA of 500743419_188-281-3-0-H5-F (SEQ ID NO:15) encodes protein claudin-5 of SEQ ID NO:16, comprising the amino acid sequence: MGSAALEILGLVLCLVGWGGLILACGLPMWQVTAFLDHNIVTAQTTWKGLWMSCVVQSTG HMQCKVYDSVLALSTEVQAARALTSAVLLAFVALFVTLAGAQCTTCVAPGPAKARVALTG GVLYLFCGLLALVPLCWTFANIVVREFYDPSVPVSQKYELGAALYIGWAATALLMVGGCLLCC 20 GAWVCTGRPDLSFPVKYSAPRRPTATGDNDKKNYV. Accordingly, it will be appreciated that all characteristics and uses of polypeptides of SEQ ID NO:16 described throughout the present application also pertain to the polypeptides encoded by the nucleic acids included in 500743419_188-281-3-0-H5-F. In addition, it will be appreciated that all characteristics and uses of the 25 polynucleotides of SEQ ID NO:15 described throughout the present application also pertain to the nucleic acids comprising the human cDNA in 500743419_188-281-3-0-H5-F. A preferred embodiment of the invention is directed toward the compositions of SEQ ID NO:15, SEQ ID NO:16, and 500743419_188-281-3-0-H5-F. Also preferred are polypeptide fragments having a biological activity as described herein and the polynucleotides encoding the fragments.

[0384] The cDNA of SEQ ID NO:15 is a polymorphic variant of the human claudin-5 protein 30 belonging to the PMP22-Claudin family of proteins characterized by four membrane-spanning segments. Claudin-5 is 218 amino acids long and contains four membrane-spanning segments. Claudin-5 contains the Claudin family signature and is encoded by a gene located on chromosome 22, specifically at position 22q11.2.

[0385] The Claudin family of proteins comprises more than twenty small glycoproteins with four predicted transmembrane domains. Claudin-5 is a component of tight junction (TJ) strands that play a role in regulation of cell permeability and polarity. Polarized epithelial and endothelial cells form barriers that separate biological compartments and regulate homeostasis. The TJ is a specialized membrane domain at the most apical region of polarized epithelial and endothelial cells that constitute continuous seals around cells. These seals serve as a physical barrier preventing solutes and water from passing freely through the paracellular space. The TJ also restricts the lateral diffusion of membrane lipids and proteins to maintain the cellular polarity. Many diseases marked by changes in cell barrier permeability are in turn related to alteration in TJ function. For example, the increase in microvascular permeability in tumors, contributing to clinically severe symptoms, appears to be the result of a dysregulation of junctional proteins. Increased TJ permeability of the epithelium, and consequently a decrease in epithelial barrier function, precedes the development of some tumors, including carcinomas and adenomas. The lung epithelium forms a barrier that allergens must cross before they can cause sensitization. The nonspecific disruption of intercellular TJs by allergens increases epithelial permeability, allowing allergens to cross the epithelial barrier. Finally, TJ permeability can also be modified by different bacterial toxins, cytokines, hormones and drugs.

[0386] An embodiment of the invention is directed to a composition comprising a claudin-5 polypeptide sequence of SEQ ID NO:16.

[0387] A further embodiment of the invention is directed to a composition comprising a claudin-5 polypeptide fragment having biological activity.

[0388] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence of SEQ ID NO:15 encoding a claudin-5 polypeptide.

[0389] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence encoding a claudin-5 polypeptide fragment having biological activity.

[0390] A further embodiment of the invention is directed to a composition comprising an antibody recognizing a claudin-5 polypeptide sequence of SEQ ID NO:16 or a claudin-5 polypeptide fragment having biological activity. Preferably, the antibody recognizes a non-linear epitopes, and binds to claudin-5 but not to claudin-5.

[0391] A preferred aspect of the invention is a host cell recombinant for polynucleotides encoding a claudin-5 polypeptide or a biologically active fragment thereof. An additional preferred aspect is a host cell recombinant for polynucleotides capable of directing claudin-5 expression. Preferably, the polynucleotides capable of directing claudin-5 expression are located in the 5'

regulatory region of the claudinyn-5 gene. Further preferably, these polynucleotides are located within 500 base pairs of the claudinyn-5 coding region. These polynucleotides preferably comprise a promoter sequence. Techniques known in the art for introducing polynucleotide sequences to endogenous sequences are described in U.S. Patent 5641670 and PCT WO9629411, which disclosures 5 are hereby incorporated by reference in their entireties. Claudinyn-5 protein produced by said host cell may be used for in vitro detection and purification methods as well as diagnosis and in vivo applications.

[0392] Another embodiment relates to a method of producing claudinyn-5 polypeptides comprising the steps of: i) transfecting a mammalian host cell with a recombinant expression vector 10 comprising a polynucleotide of the present invention, and ii) purifying the produced protein. The purification of the protein can be done following any technique well-known to those skilled in the art. Preferably, an antibody directed against claudinyn-5 or part thereof may be bound to a chromatographic support to form an affinity chromatography column. Even more preferably, the antibodies bind to claudinyn-5 but not to claudin-5.

[0393] In another embodiment, the present invention provides for a method to increase the intestinal absorption of hydrophilic drugs by increasing the paracellular permeability for said drugs either *in vitro* or *in vivo*. Indeed, the intestinal epithelium is a major barrier to the absorption of hydrophilic drugs. The presence of intercellular junctional complexes, and particularly the TJs, renders the epithelium impervious to hydrophilic drugs, which cannot diffuse across the cells through 15 the lipid bilayer of the cell membranes. The expression of claudinyn-5 can be inhibited or decreased for example using antisense-polynucleotides corresponding to claudinyn-5 to down-regulate expression of said proteins. Methods of designing, synthesizing, and using said antisense polynucleotides are well known to those skilled in the art and are discussed herein. The function of claudinyn-5 in TJ complexes can also be more specifically inhibited using direct or indirect inhibitor 20 molecules or antagonistic antibodies directed against the present protein. Any compound that inhibits or significantly decreases the activity of claudinyn-5 polypeptides can be identified by screening for test substances that decrease claudinyn-5 activity comprising the steps of contacting a cell with a test substance and comparing claudinyn-5 activity in the cell after exposure to the test substance to that of an unexposed control cell. Preferably, claudinyn-5 activity is determined by measuring the 25 permeability of the TJs using, for example, methods discussed in U.S. Patent 6,110,747, which disclosures are hereby incorporated by reference in its entirety. Preferably, such methods are used is to temporally increase epithelial permeability for therapeutic ends including but not limited to 30

increasing the efficiency of drug delivery. More preferably, for *in vivo* purposes, formulations are those comprising said compositions that are compatible with oral delivery.

[0394] In a further embodiment, a claudinyn-5 polypeptide-specific antibody or a fragment thereof can be used as a component of drug delivery vehicles such as colloids or liposomes to 5 specifically target therapeutic agents to epithelial cells. Such methods comprise the step of incorporating a claudinyn-5 polypeptide-specific antibody or a fragment thereof, into liposomes used to specifically target therapeutic agents carried by the liposomes to epithelial cells. Such drug delivery systems can be designed using any techniques known to those skilled in the art. In one such embodiment, claudinyn-5-specific antibodies or fragments thereof can also be directly conjugated 10 (either recombinantly or by using chemically active agents) to chemotherapeutic agents, radioisotopes, or prodrugs and said conjugate subsequently used in therapeutic regimens.

[0395] The negative effects of the usual preservation solutions on epithelial and endothelial permeability in organs to be transplanted are generally known, considering the disorganization of TJ proteins. This is responsible for the observed tissue injury and edema. In another embodiment, 15 agonists of the expression of claudinyn-5 can be used to maintain the content and integrity of TJs in organs to be transplanted. Further preferred embodiment is a method of using such agonists as constituents of the preservation solutions in order to improve the quality of said organs to be transplanted. Any compound that stimulates or significantly increases the expression of claudinyn-5 polypeptides can be used in one such embodiment. Such compounds can be identified by screening 20 for test substances that increase claudinyn-5 expression comprising the steps of contacting a cell with a test substance and comparing claudinyn-5 expression in the cell after exposure to the test substance to that of an unexposed control cell. In this embodiment, the activity of claudinyn-5 polypeptides can also be increased to preserve TJ function by administering a compound that causes an increase in the 25 activity of the protein. Such compounds can be identified by screening for test substances that increase claudinyn-5 activity comprising the steps of contacting a cell with a test substance and comparing claudinyn-5 activity in the cell after exposure to the test substance to that of an unexposed control cell. Preferably, claudinyn-5 activity is determined by measuring the permeability of the TJs using methods as mentioned above.

[0396] The present invention also provides animal models generated by modulating the 30 expression of the present protein in one or more tissues of the animal. Such animals are useful for a number of purposes, because they represent an *in vivo* assay method for testing candidate molecules potentially useful for the treatment of various pathophysiological aspects of diseases specifically related to the function of claudinyn-5. Study of the phenotype of such models can also allow the

identification of additional human equivalent diseases caused by or linked with claudinyn-5 deficiency. These animals can be generated with any method of targeting overexpression or inactivation of claudinyn-5. Such models are extremely useful, e.g. in the assessment of candidate therapies and drugs for the treatment of inflammatory diseases and conditions.

5 [0397] In other embodiment, the current invention is used to diagnose diseases or disorders associated with altered expression of claudinyn-5. In particular, it is useful in diagnosing patients with deficient amounts of claudinyn-5 which results in uncontrolled cell permeability. Examples of such diseases and disorders include, but are not limited to, Crohn's disease, ulcerative colitis, irritable bowel syndrome, allergic asthma, gliobastoma, colorectal carcinoma, and adenomatous polyps. The
10 method includes the steps of i) contacting a biological sample with a compound capable of selectively binding claudinyn-5 polypeptide or nucleic acids, and ii) detecting the level, or any other detectable property of claudinyn-5 in the sample. Preferably, a difference in the level or other property in the sample relative to in a control sample indicates the presence of the disease or disorder, or of a propensity for developing the disease or disorder. Preferably, the biological sample is obtained from
15 an individual suspected of suffering from the disease or condition, or at risk of developing the disease or condition. Preferably, the biological sample comprises an epithelial cell sample. A polyclonal or monoclonal antibody or any immunologically active fragment specific for claudinyn-5 may be used to detect levels of the protein. Preferably, the antibody or fragment thereof is detectably labeled with, for example, a fluorescent, radioactive, or otherwise detectable compound. Detection can be carried out
20 directly or indirectly with known immunohistological and immunofluorescing processes. For this, the reagents contained in the kit according to the invention can be directly labeled with generally known molecules, including, but not limited to, enzymes such as alkaline phosphatase and peroxidase and fluorescent dyes such as FITC, rhodamine, and Texas-Red. However, labeling can also occur indirectly by using secondary antibodies labeled with molecules such as biotin, digoxigenin or the like
25 and are then detected with a secondary reagent. A method using a nucleic acid probe corresponding to claudinyn-5 can also be used to determine the expression of claudinyn-5 using, for example, well known PCR or RT-PCR techniques and in particular in with "real-time" PCR system.

[0398] Alternatively, the present invention provides a tool to correlate modulations in the expression of claudinyn-5 with certain pathologies. Thus, the present invention provides a novel
30 candidate gene for such conditions. Also in this embodiment, the determination of the level of the expression of claudinyn-5 within a biological sample provides an assay for an *in vivo* marker to classify and/or characterize of disease states. Preferably, such a method is applied to diagnose or classify pathologies such as Crohn's disease, ulcerative colitis, irritable bowel syndrome, allergic

asthma, gliobastoma, colorectal carcinoma, and adenomatous polyps. The invention thus includes test kits useful for the quantification of the amount of claudinyn-5 in a biological sample. The kits comprise at least one immunological binding partner, e.g. a monoclonal or polyclonal antibody specific for claudinyn-5 and coupled to detectable markers. Application of such assays can be used to
5 monitor the progress of therapy administered to treat these or other conditions. Thus the assays may be applied in any situation wherein the level of claudinyn-5 expression can be used as an index of the condition, treatment, or effect of substances directly administered to the subject or to which the subject is exposed in the environment. Thus, the condition of a patient can be monitored continuously and the quantified amount of claudinyn-5 polypeptides measured in the pathological sample can be compared
10 with the amount quantified in a biological sample of a normal individual or with the previous analysis of the same patient. In this embodiment, this marker can be measured by any suitable method, including immunoassays. It can also preferably be measured in tissues and fluids recovered from inflammatory sites.

[0399] A further embodiment of the present invention is to provide novel methods and
15 compositions useful for the treatment of diseases and conditions associated with impaired TJ function. More preferably, claudinyn-5 polypeptide or fragment thereof can be used to treat diseases associated with an impaired claudinyn-5 related function, including but not limited to, Crohn's disease, ulcerative colitis, irritable bowel syndrome, allergic asthma, gliobastoma, colorectal carcinoma, and adenomatous polyps. Such methods comprise the administration of a therapeutically-effective amount of claudinyn-
20 5 to mammals suffering from the disease or condition, where "effective amount" is meant a concentration of claudinyn-5 capable of significantly restore a physiological TJs function. The compositions may be administrated using any of the methods known in the art or discussed herein for delivering claudinyn-5 polynucleotides to a cell. Any agonists of the expression or the activity of claudinyn-5 polypeptides identified as mentioned above can also be used in this embodiment. Such
25 compositions are preferably delivered to an individual in combination with a physiologically acceptable carrier, such as a saline solution or other physiologically buffer suitable for administration to a patient. The particular amount of the compositions of the invention that will be administered to the mammal for any particular condition will depend on the clinical condition of the patient, and other factors such as the weight, age, and route of delivery. These compositions can optionally comprise,
30 one or more other compound of interest. This co-administration may be by simultaneous administration or by separate or sequential administrations. All of these additional components may be either obtained from natural sources or produced by recombinant genetic engineering techniques and/or chemical modification.

[0400] In a further embodiment, the present invention provides a method of producing “bioartificial” epithelia from non-epithelial cells. The “bioartificial” epithelia produced according to the invention provides various clinical applications for the generation of epithelia unable to repair or regenerate themselves. It can be used, for example, for reconstructive surgical procedures, for treating 5 of disorders related to epithelial loss (for hereditary, traumatic or oncological reasons) or for another therapeutic purposes (for example, burn treatments). Further preferred is a method of using the nucleic acid corresponding to claudinyn-5 wherein “bioartificial” epithelial cells are obtained by transfection of said nucleic acid, and consequential remodeling, of the patient’s autologous cells not affected by any of the above conditions. The use of autologous cells in the preparation of the “bioartificial” 10 epithelial cells of the invention in methods of treating disorders, conditions, or diseases associated with the loss of epithelial cells reduces or eliminates the risk of tissue rejection typically observed in transplantation methodologies. Methods of bioartificial tissue engineering are generally known to those skilled in the art (Machens H.G. et al., Cells Tissues Organs 167: 88-94 (2000), the disclosure of which is hereby incorporated by reference in its entirety).

15 Protein of SEQ ID NO:18 (internal designation Clone 645730_181-16-1-0-G9-F)

[0401] The cDNA of Clone 645730_181-16-1-0-G9-F (SEQ ID NO:17) encodes Benzodiazepine Receptor 3 (BZRP-R3) protein of SEQ ID NO:18, comprising the amino acid sequence:

MRLQGAIJVLLPHLGPILVWLFRDHMSGWCEGPRMLSWC¹CPFYKV²LLL³VQTAIYSVVG⁵YASY
20 LVWKDLGGGLGWPLALPLGLYADQLTISWTVLVLFFT⁷VHN⁸PGLALLH⁹LLL¹⁰YGLVVSTALIW¹¹
HPINKLA¹²ALLLPYLAWLTVTSALTYHLWRDSLCPVHQ¹⁴QPQ¹⁵PKTEKSD. Accordingly, it will be appreciated that all characteristics and uses of the polypeptides of SEQ ID NO:18 described throughout the present application also pertain to the polypeptides encoded by the nucleic acids included in Clone 645730_181-16-1-0-G9-F. In addition, it will be appreciated that all characteristics 25 and uses of the polynucleotides of SEQ ID NO:17 described throughout the present application also pertain to the nucleic acids comprising the human cDNA in Clone 645730_181-16-1-0-G9-F. A preferred embodiment of the invention is directed toward the compositions of SEQ ID NO:17, SEQ ID NO:18 and Clone 645730_181-16-1-0-G9-F. Also preferred are polypeptide fragments having a biological activity as described herein and the polynucleotides encoding the fragments.

30 [0402] The protein of SEQ ID NO:18, BZRP-R3, is a novel polymorphic variant of human peripheral benzodiazepine receptor/isoquinoline binding protein (PBR/IBP) (accession numbers Q9Y531). BZRP-R3 displays five transmembrane-spanning segments: LQGAIFVLLPHLGPILVWLFT, VLLLVQTAIYSVVG⁵YASYLVW,

GLYADQLTISWTVLVLFFTVH, GLALLHLLLLYGLVVSTALIW, and LAALLLLPYLAWLTTSALTY. Accordingly, some embodiments of the present invention relate to polypeptides comprising the transmembrane domains. Moreover, BZRP-R3 displays a stretch of 11 amino acids (VTSALTYHLWR) that bind cholesterol. Accordingly, a preferred embodiment of the 5 present invention comprises the amino acids of the cholesterol recognition/interaction domain and the polynucleotides encoding the same.

[0403] BZRP-R3 is capable of binding benzodiazepine and imidazopyridine derivatives, but is distinct from the GABA neurotransmitter receptor. BZRP-R3 polypeptides are most abundant in steroidogenic cells and are found primarily on outer mitochondrial membranes. BZRP-R3 is 10 associated with a 34-kDa pore-forming, voltage-dependent anion channel protein located on the outer/inner mitochondrial membrane contact sites. Ligands of BZRP-R3, upon binding to the receptor, simulate steroid synthesis in steroidogenic cells in vitro and in vivo. BZRP-R3 stimulates steroid formation by increasing the rate of cholesterol transfer from the outer to the inner mitochondrial membrane.

15 [0404] In addition to its role in mediating cholesterol movement across membranes, BZRP-R3 has been implicated in several other physiological functions, including cell growth and differentiation, chemotaxis, mitochondrial physiology, porphyrin and heme biosynthesis, immune response, and anion transport. In addition, BZRP-R3 agonists are potent anti-apoptotic compounds.

[0405] BZRP-R3 is associated with stress and anxiety disorders. BZRP-R3 plays a role in the 20 regulation of several stress systems such as the HPA axis, the sympathetic nervous system, the renin-angiotensin axis, and the neuroendocrine axis. In these systems, acute stress typically leads to increases in BZRP-R3 density, whereas chronic stress typically leads to decreases in BZRP-R3 density. For example, in Generalized Anxiety Disorder (GAD), Panic Disorder (PD), Generalized Social Phobia (GSP), and Post-Traumatic Stress Disorders (PTSD), BZRP-R3 density is typically 25 decreased. BZRP-R3 is expressed glial cells in the brain. Furthermore, BZRP-R3 expression is increased in neurodegenerative disorders and after neurotoxic and traumatic-ischemic brain damage. BZRP-R3 expression is decreased in chronic schizophrenics, suggesting that the decreased density of BZRP-R3 in the brain may be involved in the pathophysiology of schizophrenia. However, BZRP-R3 is higher than normal in autopsied brain tissue from PSE patients (Portal-Systemic Encephalopathy 30 patients).

[0406] BZRP-R3 increases mitochondrial activity and prevents apoptosis and is therefore implicated in tumor cell proliferation. BZRP-R3 is preferentially expressed in liver and breast cancers.

Further, BZRP-R3 is useful as a tool/marker for detection, diagnosis, prognosis and treatment of cancer.

[0407] Many ligands have been described that bind to BZRP-R3 with various affinities. Some benzodiazepines, Ro 5-4864 [4-chlorodiazepam], diazepam and structurally related compounds, 5 are potent and selective PBR ligands. Exogenous ligands also include 2-phenylquinoline carboxamides (PK11195 series), imidazo [1,2-a]pyridine-3-acetamides (Alpidem series), pyridazine, and isoquinilone derivatives. Some endogenous compounds, including porphyrins and diazepam binding inhibitor (DBI), bind to BZRP-R3.

[0408] A preferred aspect of the invention is a host cell recombinant for polynucleotides 10 encoding a BZRP-R3 polypeptide or a biologically active fragment thereof. An additional preferred aspect is a host cell recombinant for polynucleotides capable of directing BZRP-R3 expression. Preferably, the polynucleotides capable of directing BZRP-R3 expression are located in the 5' regulatory region of the BZRP-R3 gene. Further preferably, these polynucleotides are located within 500 base pairs of the BZRP-R3 coding region. These polynucleotides preferably comprise a promoter 15 sequence. Techniques known in the art for introducing polynucleotide sequences to endogenous sequences are described in U.S. Patent 5641670 and PCT WO9629411, which disclosures are hereby incorporated by reference in their entireties. BZRP-R3 protein produced by said host cell may be used for in vitro detection and purification methods as well as diagnosis and in vivo applications.

[0409] A preferred embodiment of the invention relates to compositions and methods using 20 the protein of the invention or fragment thereof to label mitochondria in order to visualize any change in number, topology or morphology of this organelle, for example in association with a mitochondria-related human disorder, such as hereditary spastic paraplegia, neuroleptic malignant syndrome (NMS), the Rett syndrome, Alpers disease, or mitochondrial encephalomyopathies. For example, the protein may be rendered easily detectable by inserting the cDNA encoding the protein of the invention into a 25 eukaryotic expression vector in frame with a sequence encoding a tag sequence. Eukaryotic cells expressing the tagged protein of the invention may also be used for the *in vitro* screening of drugs or genes capable of treating any mitochondria-related disease or conditions.

[0410] A preferred embodiment of the invention is a method to detect mitochondria comprising the step of contacting an antibody specific for BZRP-R3 with a cell. Preferably, the cell is 30 fixed or otherwise permeable to antibodies. Preferably, the antibody is labeled with any detectable moiety including, but are not limited to, a fluorescent label, a radioactive atom, a paramagnetic ion, biotin, a chemiluminescent label or a label which can be detected through a secondary enzymatic or

binding step. The invention further provides a method of diagnosing mitochondria-related disorders or conditions, such as those listed above, and distinguishing these from other disorders.

[0411] In another embodiment, the protein of the invention may be used to target heterologous compounds (polynucleotides) to steroidogenic cells and/or the mitochondria. Such recombinant cDNA may be introduced, for example, using in any vector, viral or non-viral, and viral vectors can be but not limited to retroviral, adenoviral, and adeno-associated vectors. For example, these heterologous polynucleotides may be used to deliver nucleic acids for mitochondrial gene therapy, i.e. to replace a defective mitochondrial gene and/or to inhibit the deleterious expression of a mitochondrial gene.

[0412] Another preferred embodiment of the invention is a method of screening for compounds that modulate the expression of BZRP-R3. This method comprises the steps of i) contacting a cell with a test compound and ii) comparing the level of BZRP-R3 polypeptides in a cell after exposure to the test compound to that of an untreated control cell. The level of BZRP-R3 polypeptides may be inferred by detecting mRNA for BZRP-R3 by methods common to the art such as Northern blotting or RT-PCR. The level of BZRP-R3 polypeptides may also be detected by antibody-based methods common to the art such as Western blotting or immunofluorescence. Test compounds that increase BZRP-R3 expression are useful as agonists, as discussed herein. Test compounds that decrease BZRP-R3 expression are useful as antagonists, as discussed herein.

[0413] Antagonists of BZRP-R3 include agents which decrease the levels of expressed mRNA encoding the protein of SEQ ID NO:18. These include, but are not limited to, RNAi, one or more ribozymes capable of digesting the protein of the invention, or antisense oligonucleotides capable of hybridizing to mRNA encoding BZRP-R3. Antisense oligonucleotides can be administrated as DNA, RNA, as DNA entrapped in proteoliposomes containing viral envelope receptor proteins [Kanoda, Y. et al. (1989) Science 243: 375, which disclosure is hereby incorporated by reference in its entirety] or as part of a vector which can be expressed in the target cell to provide antisense DNA or RNA. Vectors which are expressed in particular cell types are known in the art. Alternatively, the DNA can be injected along with a carrier. A carrier can be a protein such as a cytokine, for example interleukin 2, or polylysine-glycoprotein carriers. Carrier proteins, vectors, and methods of making and using polylysine carrier systems are known in the art. Alternatively, nucleic acids encoding antisense molecules may be coated onto gold beads and introduced into the skin with, for example, a gene gun [Ulmer, J.B. et al. (1993) Science 259:1745, which disclosure is hereby incorporated by reference in its entirety].

[0414] A preferred embodiment of the invention is a method of screening for compounds that bind to BZRP-R3 polypeptides. Such compounds are useful for developing agonists and antagonists of BZRP-R3 activity. This method comprises the steps of: i) contacting a BZRP-R3 polypeptide or fragment thereof with a test compound under conditions that allow binding to occur and ii) detecting binding of said test compound. Binding may be detected by any method common to the art such as competition with a labeled antibody specific for BZRP-R3 or by direct labeling of each test substance. In one example of such a method, a polynucleotide encoding a BZRP-R3 polypeptide or a biologically active fragment thereof is transformed into a eukaryotic or prokaryotic host cell. The transformed cells may be viable or fixed. Drugs or compounds which are candidates for binding BZRP-R3 polypeptides are screened against such transformed cells in binding assays well known to those skilled in the art. Alternatively, assays such as those taught in Geysen H. N., WO Application 84/03564, published on Sep. 13, 1984, and incorporated herein by reference in its entirety, may be used to screen for peptide compounds which demonstrate binding affinity for BZRP-R3 polypeptides or fragments thereof. In another embodiment, competitive drug screening assays using neutralizing antibodies specifically compete with a test compound for binding to BZRP-R3 polypeptides or fragments thereof. Preferred test compounds are those included in the benzodiazepine class, such as diazepam (i.e., valium), triazolobenzodiazepine, and adinazolam, as well as modified versions thereof. Further preferred test compounds are in the imidazopyridine and isoquinilone classes. Each test compound may additionally be excluded individually or as a class.

[0415] A variety of drug screening techniques may be employed. In this aspect of the invention, BZRP-R3 polypeptide or fragments thereof, may be free in solution, affixed to a solid support, recombinantly expressed on or chemically attached to a cell surface, or located intracellularly. The formation of binding complexes between BZRP-R3 polypeptides or fragments thereof, and the compound being tested, may then be measured as described.

[0416] Another embodiment of the subject invention provides compositions and methods of selectively modulating the activity of the protein of the invention. Modulation of BZRP-R3 allows for the successful prevention, treatment, or management of disorders or biochemical abnormalities associated with BZRP-R3. Agonist compounds are those that increase the amount of BZRP-R3 polypeptides in a cell or increase the biological activity of BZRP-R3. A preferred embodiment of the invention is a method of screening for agonists that bind to BZRP-R3 comprising the steps of: i) screening for test substances that bind to BZRP-R3, as described above and ii) detecting BZRP-R3 biological activity. Preferably, this method is accomplished in an intact cell. Further preferably, the cell is a steroidogenic cell such as a testicular or ovarian cell. Preferably, the biological activity of

BZRP-R3 is determined by measuring the concentration of steroid hormones released from the cell before and after exposure to the test substance. Agonists of BZRP-R3 will increase the release of steroid hormones from the cell. Antagonist compounds are those that decrease the amount of BZRP-R3 polypeptides in a cell or decrease the biological activity of BZRP-R3. Another preferred embodiment of the invention is a method of screening for antagonists that bind to BZRP-R3 comprising the steps of: i) screening for test substances that bind to BZRP-R3, as described above and ii) detecting BZRP-R3 biological activity. Preferably, this method is accomplished in an intact cell. Further preferably, the cell is a steroidogenic cell such as a testicular or ovarian cell. Preferably, the biological activity of BZRP-R3 is determined by measuring the concentration of steroid hormones released from the cell before and after exposure to the test substance. Antagonists of BZRP-R3 will decrease the release of steroid hormones from the cell.

[0417] Antagonists, able to reduce or inhibit the expression or the activity of the protein of the invention, are useful in the treatment of diseases associated with elevated levels of BZRP-R3, increased cell proliferation or reduced apoptosis, and increased cholesterol transport. Thus, the subject invention provides methods for treating a variety of diseases or disorders, including but not limited to cancers, especially liver and breast cancer, and portal-systemic encephalopathy. Increased cholesterol transport into the mitochondria of steroidogenic cells results in higher than normal production of steroid hormones such as progesterone, testosterone, and estrogen. Abnormally high levels of steroid hormones lead to disruption of adrenocortical feedback mechanisms and underproduction of trophic hormones from the hypothalamus and pituitary. Inhibition of BZRP-R3 and steroidogenesis may increase levels of trophic hormones such as gonadotropin-releasing hormone.

[0418] Alternatively, the subject invention provides a method of treating diseases or disorders associated with decreased levels of BZRP-R3 polypeptides and decreased steroid hormone release with an agonist thereof. Such method comprises the step of contacting a cell with a BZRP-R3 agonist. This method comprises the step of contacting a cell with an agonist of BZRP-R3. Thus, the subject invention provides methods of treating disorders including, but not limited to, schizophrenia, chronic stress, GAD, PD, GSP and PTSD. Other disorders which may be treated by agonists of BZRP-R3 include those associated with decreases in cell proliferation, e.g. developmental retardation. Furthermore, because BZRP-R3 is able to transport cholesterol into cells, BZRP-R3 agonists may also be used to increase cholesterol transport into cells. Diseases associated with cholesterol transport deficiencies include lipoidal adrenal hyperplasia, ovarian cysts, abnormal lipid deposits in steroidogenic cells. Disorders that reflect a requirement for cholesterol for myelin and myelination, include Alzheimer's disease, multiple sclerosis, spinal cord injury, and brain development neuropathy.

The methods of treating disorders associated with decreased levels of BZRP-R3 may be practiced by introducing agonists which stimulate the expression or the activity of BZRP-R3.

[0419] Additionally, disorders resulting from defective mitochondrial activity may be treated with an agonist to BZRP-R3. Defective mitochondrial activity may alternatively or additionally result 5 in the generation of highly reactive free radicals that have the potential of damaging cells and tissues. These free radicals may include reactive oxygen species (ROS) such as superoxide, peroxynitrite and hydroxyl radicals, and other reactive species that may be toxic to cells and cause apoptosis. For example, oxygen free radical induced lipid peroxidation is a well-established pathogenic mechanism in central nervous system (CNS) injury such as that found in a number of degenerative diseases, and in 10 ischemia (i.e., stroke). Diseases associated with altered mitochondrial function and apoptosis include: Alzheimer's Disease, diabetes mellitus, Parkinson's Disease, Huntington's disease, dystonia, Leber's hereditary optic neuropathy, schizophrenia, mitochondrial encephalopathy, lactic acidosis, and stroke.

[0420] A further preferred embodiment includes a method of inhibiting apoptosis of cells in culture. This method comprises the step of contacting a cell in culture with an agonist to BZRP-R3. 15 Such methods are useful for culturing cells that are notoriously undergo apoptosis, such as primary neurons and lymphocytes.

[0421] In one embodiment, the level of BZRP-R3 in a cell may be increased by introducing nucleic acids encoding a BZRP-R3 polypeptide or biologically active fragment thereof into a targeted cell type. Vectors useful in such methods are known to those skilled in the art, as are methods of 20 introducing such nucleic acids into target tissues.

[0422] Antibodies or other polypeptides capable of reducing or inhibiting the activity of BZRP-R3 may be provided as in isolated and substantially purified form. Alternatively, antibodies or other polypeptides capable of inhibiting or reducing the activity of BZRP-R3 may be recombinantly expressed in the target cell to provide a modulating effect. In addition, compounds which inhibit or 25 reduce the activity of BZRP-R3 may be incorporated into biodegradable polymers being implanted in the vicinity of where drug delivery is desired. For example, biodegradable polymers may be implanted at the site of a tumor or, alternatively, biodegradable polymers containing antagonists/agonists may be implanted to slowly release the compounds systemically. Biodegradable polymers, and their use, are known to those of skill in the art (see, for example, Brem et al. (1991) J. Neurosurg. 74:441-446, 30 which disclosure is hereby incorporated by reference in its entirety).

[0423] In another embodiment, the invention provides methods and compositions for detecting the level of expression of the mRNA encoding the protein of the invention. Quantification of

mRNA levels of BZRP-R3 may be useful for the diagnosis or prognosis of diseases associated with an altered expression of the protein of the invention. Assays for the detection and quantification of the mRNA encoding BZRP-R3 are well known in the art (see, for example, Maniatis, Fitzsch and Sambrook, Molecular Cloning; A Laboratory Manual (1982), or Current Protocols in Molecular 5 Biology, Ausubel, F.M. et al. (Eds), Wiley & Sons, Inc., disclosures of which are hereby incorporated by reference in their entireties).

[0424] Polynucleotides probes or primers for the detection of BZRP-R3 mRNA can be designed from the cDNA of SEQ ID NO:17. Methods for designing probes and primers are known in the art. In another embodiment, the subject invention provides diagnostic kits for the detection of the 10 mRNA of the protein of the invention in cells. The kit comprises a package having one or more containers of oligonucleotide primers for detection of the polynucleotides of the invention in PCR assays or one or more containers of polynucleotide probes for the detection of the BZRP-R3 mRNA by in situ hybridization or Northern analysis. Kits may, optionally, include containers of various reagents used in various hybridization assays. The kit may also, optionally, contain one or more of the 15 following items: polymerization enzymes, buffers, instructions, controls, or detection labels. Kits may also, optionally, include containers of reagents mixed together in suitable proportions for performing the hybridization assay methods in accordance with the invention. Reagent containers preferably contain reagents in unit quantities that obviate measuring steps when performing the subject methods.

[0425] In another embodiment, the invention relates to methods and compositions for 20 detecting and quantifying the level of the protein of the invention present in a particular biological sample. These methods are useful for the diagnosis or prognosis of diseases associated with altered levels of the protein of the invention. Diagnostic assays to detect the protein of the invention may comprise a biopsy, in situ assay of cells from organ or tissue sections, or an aspirate of cells from a tumor or normal tissue. In addition, assays may be conducted upon cellular extracts from organs, 25 tissues, cells, urine, or serum or blood or any other body fluid or extract.

[0426] Assays for the quantification of BZRP-R3 polypeptides may be performed according to methods well known in the art. Typically, these assays comprise the steps of: contacting the sample with a ligand of the protein of the invention or an antibody (polyclonal or monoclonal) that specifically recognizes the protein of the invention or a fragment thereof and detecting the complex formed 30 between the protein of the invention present in the sample and the ligand or antibody. Fragments of the ligands and antibodies may also be used in the binding assays, provided these fragments are capable of specifically interacting with BZRP-R3 polypeptides. Further, ligands and antibodies which bind to BZRP-R3 may be labeled according to methods known in the art. Labels which are useful in

the subject invention include, but are not limited to, enzymes labels, radioisotopic labels, paramagnetic labels, and chemiluminescent labels. Common techniques are described by Kennedy, J. H., et al. (1976) Clin. Chim. Acta 70:1-31; and Schurs, A. H. et al. (1977) Clin. Chim. Acta 81: 1-40, disclosures of which are hereby incorporated by reference in their entireties.

5 [0427] The subject invention also provides methods and compositions for the identification of metastatic tumor masses. In this aspect of the invention, the polypeptide or antibody that specifically binds a BZRP-R3 polypeptide or fragment thereof may be used as a marker for the identification of the metastatic tumor mass. Metastatic tumors originating from the breast or liver may overexpress BZRP-
R3 polypeptides, whereas newly forming tumors, or those originating from other tissues are not
10 expected to bear BZRP-R3.

Protein of SEQ ID NO:20 (internal designation Clone 646762_181-21-2-0-A3-F)

[0428] The cDNA of Clone 646762_181-21-2-0-A3-F (SEQ ID NO:19) encodes Benzodiazepine Receptor 4 (BZRP-R4) protein of SEQ ID NO:20, comprising the amino acid sequence:

15 MRLQGAIFVLLPHLGPILVWLFRDHMSGLCEGPRMLSWCPTYKVLLVQTAIYSVVGYASY
LVWKDLGGGLGWPLALPLGLYAVQLTISWTVLVLFFTvhNPGLALLHLLLGYGLVVSTALIW
HPINKLAALLLPYLAWLTVTSALTYHLWRDSLCPVHQQPQPTEKSD. Accordingly, it will be
appreciated that all characteristics and uses of the polypeptides of SEQ ID NO:20 described
throughout the present application also pertain to the polypeptides encoded by the nucleic acids
20 included in Clone 646762_181-21-2-0-A3-F. In addition, it will be appreciated that all characteristics
and uses of the polynucleotides of SEQ ID NO:19 described throughout the present application also
pertain to the nucleic acids comprising the human cDNA in Clone 646762_181-21-2-0-A3-F. A
preferred embodiment of the invention is directed toward the compositions of SEQ ID NO:19, SEQ ID
NO:20 and Clone 646762_181-21-2-0-A3-F. Also preferred are polypeptide fragments having a
25 biological activity as described herein and the polynucleotides encoding the fragments.

[0429] The protein of SEQ ID NO:20, BZRP-R4, is a novel polymorphic variant of human peripheral benzodiazepine receptor/isoquinoline binding protein (PBR/IBP) (accession numbers Q9Y531). BZRP-R4 displays five transmembrane-spanning segments:
LQGAIFVLLPHLGPILVWLFT, VLLLVQTAIYSVVGYASYLVW,
30 LYAVQLTISWTVLVLFFTvhN, LALLHLLLGYGLVVSTALIW, and
LAALLLPYLAWLTVTSALTY. Accordingly, some embodiments of the present invention relate to polypeptides comprising the transmembrane domains. Moreover, BZRP-R4 displays a stretch of 11

amino acids (VTSALTYHLWR) that bind cholesterol. Accordingly, a preferred embodiment of the present invention comprises the amino acids of the cholesterol recognition/interaction domain and the polynucleotides encoding the same.

[0430] BZRP-R4 is capable of binding benzodiazepine and imidazopyridine derivatives, but 5 is distinct from the GABA neurotransmitter receptor. BZRP-R4 polypeptides are most abundant in steroidogenic cells and are found primarily on outer mitochondrial membranes. BZRP-R4 is associated with a 34-kDa pore-forming, voltage-dependent anion channel protein located on the outer/inner mitochondrial membrane contact sites. Ligands of BZRP-R4, upon binding to the receptor, simulate steroid synthesis in steroidogenic cells in vitro and in vivo. BZRP-R4 stimulates steroid 10 formation by increasing the rate of cholesterol transfer from the outer to the inner mitochondrial membrane.

[0431] In addition to its role in mediating cholesterol movement across membranes, BZRP- 15 R4 has been implicated in several other physiological functions, including cell growth and differentiation, chemotaxis, mitochondrial physiology, porphyrin and heme biosynthesis, immune response, and anion transport. In addition, BZRP-R4 agonists are potent anti-apoptotic compounds.

[0432] BZRP-R4 is associated with stress and anxiety disorders. BZRP-R4 plays a role in the regulation of several stress systems such as the HPA axis, the sympathetic nervous system, the renin-angiotensin axis, and the neuroendocrine axis. In these systems, acute stress typically leads to increases in BZRP-R4 density, whereas chronic stress typically leads to decreases in BZRP-R4 20 density. For example, in Generalized Anxiety Disorder (GAD), Panic Disorder (PD), Generalized Social Phobia (GSP), and Post-Traumatic Stress Disorders (PTSD), BZRP-R4 density is typically decreased. BZRP-R4 is expressed glial cells in the brain. Furthermore, BZRP-R4 expression is increased in neurodegenerative disorders and after neurotoxic and traumatic-ischemic brain damage. BZRP-R4 expression is decreased in chronic schizophrenics, suggesting that the decreased density of 25 BZRP-R4 in the brain may be involved in the pathophysiology of schizophrenia. However, BZRP-R4 is higher than normal in autopsied brain tissue from PSE patients (Portal-Systemic Encephalopathy patients).

[0433] BZRP-R4 increases mitochondrial activity and prevents apoptosis and is therefore implicated in tumor cell proliferation. BZRP-R4 is preferentially expressed in liver and breast cancers. 30 Further, BZRP-R4 is useful as a tool/marker for detection, diagnosis, prognosis and treatment of cancer.

[0434] Many ligands have been described that bind to BZRP-R4 with various affinities. Some benzodiazepines, Ro 5-4864 [4-chlorodiazepam], diazepam and structurally related compounds, are potent and selective PBR ligands. Exogenous ligands also include 2-phenylquinoline carboxamides (PK11195 series), imidazo [1,2-a]pyridine-3-acetamides (Alpidem series), pyridazine, and 5 isoquinilone derivatives. Some endogenous compounds, including porphyrins and diazepam binding inhibitor (DBI), bind to BZRP-R4.

[0435] A preferred aspect of the invention is a host cell recombinant for polynucleotides encoding a BZRP-R4 polypeptide or a biologically active fragment thereof. An additional preferred aspect is a host cell recombinant for polynucleotides capable of directing BZRP-R4 expression. 10 Preferably, the polynucleotides capable of directing BZRP-R4 expression are located in the 5' regulatory region of the BZRP-R4 gene. Further preferably, these polynucleotides are located within 500 base pairs of the BZRP-R4 coding region. These polynucleotides preferably comprise a promoter sequence. Techniques known in the art for introducing polynucleotide sequences to endogenous sequences are described in U.S. Patent 5641670 and PCT WO9629411, which disclosures are hereby 15 incorporated by reference in their entireties. BZRP-R4 protein produced by said host cell may be used for in vitro detection and purification methods as well as diagnosis and in vivo applications.

[0436] A preferred embodiment of the invention relates to compositions and methods using the protein of the invention or fragment thereof to label mitochondria in order to visualize any change in number, topology or morphology of this organelle, for example in association with a mitochondria-related human disorder, such as hereditary spastic paraplegia, neuroleptic malignant syndrome (NMS), 20 the Rett syndrome, Alpers disease, or mitochondrial encephalomyopathies. For example, the protein may be rendered easily detectable by inserting the cDNA encoding the protein of the invention into a eukaryotic expression vector in frame with a sequence encoding a tag sequence. Eukaryotic cells expressing the tagged protein of the invention may also be used for the *in vitro* screening of drugs or 25 genes capable of treating any mitochondria-related disease or conditions.

[0437] A preferred embodiment of the invention is a method to detect mitochondria comprising the step of contacting an antibody specific for BZRP-R4 with a cell. Preferably, the cell is fixed or otherwise permeable to antibodies. Preferably, the antibody is labeled with any detectable moiety including, but are not limited to, a fluorescent label, a radioactive atom, a paramagnetic ion, 30 biotin, a chemiluminescent label or a label which can be detected through a secondary enzymatic or binding step. The invention further provides a method of diagnosing mitochondria-related disease or conditions, for example, but not limited to, those listed above, and distinguishing such diseases from other diseases.

[0438] In another embodiment, the protein of the invention may be used to target heterologous compounds (polynucleotides) to steroidogenic cells and/or the mitochondria. Such recombinant cDNA may be introduced, for example, using in any vector, viral or non-viral, and viral vectors can be but not limited to retroviral, adenoviral, and adeno-associated vectors. For example, 5 these heterologous polynucleotides may be used to deliver nucleic acids for mitochondrial gene therapy, i.e. to replace a defective mitochondrial gene and/or to inhibit the deleterious expression of a mitochondrial gene.

[0439] Another preferred embodiment of the invention is a method of screening for compounds that modulate the expression of BZRP-R4. This method comprises the steps of i) 10 contacting a cell with a test compound and ii) comparing the level of BZRP-R4 polypeptides in a cell after exposure to the test compound to that of an untreated control cell. The level of BZRP-R4 polypeptides may be inferred by detecting mRNA for BZRP-R4 by methods common to the art such as Northern blotting or RT-PCR. The level of BZRP-R4 polypeptides may also be detected by antibody-based methods common to the art such as Western blotting or immunofluorescence. Test compounds 15 that increase BZRP-R4 expression are useful as agonists, as discussed herein. Test compounds that decrease BZRP-R4 expression are useful as antagonists, as discussed herein.

[0440] Antagonists of BZRP-R4 include agents which decrease the levels of expressed mRNA encoding the protein of SEQ ID NO:20. These include, but are not limited to, RNAi, one or more ribozymes capable of digesting the protein of the invention, or antisense oligonucleotides 20 capable of hybridizing to mRNA encoding BZRP-R4. Antisense oligonucleotides can be administrated as DNA, RNA, as DNA entrapped in proteoliposomes containing viral envelope receptor proteins [Kanoda, Y. et al. (1989) Science 243: 375, which disclosure is hereby incorporated by reference in its entirety] or as part of a vector which can be expressed in the target cell to provide antisense DNA or RNA. Vectors which are expressed in particular cell types are known in the art. 25 Alternatively, the DNA can be injected along with a carrier. A carrier can be a protein such as a cytokine, for example interleukin 2, or polylysine-glycoprotein carriers. Carrier proteins, vectors, and methods of making and using polylysine carrier systems are known in the art. Alternatively, nucleic acids encoding antisense molecules may be coated onto gold beads and introduced into the skin with, for example, a gene gun [Ulmer, J.B. et al. (1993) Science 259:1745, which disclosure is hereby 30 incorporated by reference in its entirety].

[0441] A preferred embodiment of the invention is a method of screening for compounds that bind to BZRP-R4 polypeptides. Such compounds are useful for developing agonists and antagonists of BZRP-R4 activity. This method comprises the steps of: i) contacting a BZRP-R4 polypeptide or

fragment thereof with a test compound under conditions that allow binding to occur and ii) detecting binding of said test compound. Binding may be detected by any method common to the art such as competition with a labeled antibody specific for BZRP-R4 or by direct labeling of each test substance. In one example of such a method, a polynucleotide encoding a BZRP-R4 polypeptide or a biologically active fragment thereof is transformed into a eukaryotic or prokaryotic host cell. The transformed cells may be viable or fixed. Drugs or compounds which are candidates for binding BZRP-R4 polypeptides are screened against such transformed cells in binding assays well known to those skilled in the art. Alternatively, assays such as those taught in Geysen H. N., WO Application 84/03564, published on Sep. 13, 1984, and incorporated herein by reference in its entirety, may be used to screen for peptide compounds which demonstrate binding affinity for BZRP-R4 polypeptides or fragments thereof. In another embodiment, competitive drug screening assays using neutralizing antibodies specifically compete with a test compound for binding to BZRP-R4 polypeptides or fragments thereof. Preferred test compounds are those included in the benzodiazepine class, such as diazepam (i.e., valium), triazolobenzodiazepine, and adinazolam, as well as modified versions thereof. Further preferred test compounds are in the imidazopyridine and isoquinilone classes. Each test compound may additionally be excluded individually or as a class.

[0442] A variety of drug screening techniques may be employed. In this aspect of the invention, BZRP-R4 polypeptide or fragments thereof, may be free in solution, affixed to a solid support, recombinantly expressed on or chemically attached to a cell surface, or located intracellularly. 20 The formation of binding complexes between BZRP-R4 polypeptides or fragments thereof, and the compound being tested, may then be measured as described.

[0443] Another embodiment of the subject invention provides compositions and methods of selectively modulating the activity of the protein of the invention. Modulation of BZRP-R4 allows for the successful prevention, treatment, or management of disorders or biochemical abnormalities 25 associated with BZRP-R4. Agonist compounds are those that increase the amount of BZRP-R4 polypeptides in a cell or increase the biological activity of BZRP-R4. A preferred embodiment of the invention is a method of screening for agonists that bind to BZRP-R4 comprising the steps of: i) screening for test substances that bind to BZRP-R4, as described above and ii) detecting BZRP-R4 biological activity. Preferably, this method is accomplished in an intact cell. Further preferably, the 30 cell is a steroidogenic cell such as a testicular or ovarian cell. Preferably, the biological activity of BZRP-R4 is determined by measuring the concentration of steroid hormones released from the cell before and after exposure to the test substance. Agonists of BZRP-R4 will increase the release of steroid hormones from the cell. Antagonist compounds are those that decrease the amount of BZRP-

R4 polypeptides in a cell or decrease the biological activity of BZRP-R4. Another preferred embodiment of the invention is a method of screening for antagonists that bind to BZRP-R4 comprising the steps of: i) screening for test substances that bind to BZRP-R4, as described above and ii) detecting BZRP-R4 biological activity. Preferably, this method is accomplished in an intact cell.

5 Further preferably, the cell is a steroidogenic cell such as a testicular or ovarian cell. Preferably, the biological activity of BZRP-R4 is determined by measuring the concentration of steroid hormones released from the cell before and after exposure to the test substance. Antagonists of BZRP-R4 will decrease the release of steroid hormones from the cell.

[0444] Antagonists, able to reduce or inhibit the expression or the activity of the protein of the 10 invention, are useful in the treatment of diseases associated with elevated levels of BZRP-R4, increased cell proliferation or reduced apoptosis, and increased cholesterol transport. Thus, the subject invention provides methods for treating a variety of diseases or disorders, including but not limited to cancers, especially liver and breast cancer, and portal-systemic encephalopathy. Increased cholesterol transport into the mitochondria of steroidogenic cells results in higher than normal production of 15 steroid hormones such as progesterone, testosterone, and estrogen. Abnormally high levels of steroid hormones lead to disruption of adrenocortical feedback mechanisms and underproduction of trophic hormones from the hypothalamus and pituitary. Inhibition of BZRP-R4 and steroidogenesis may increase levels of trophic hormones such as gonadotropin-releasing hormone.

[0445] Alternatively, the subject invention provides a method of treating diseases or disorders 20 associated with decreased levels of BZRP-R4 polypeptides and decreased steroid hormone release with an agonist thereof. Such method comprises the step of contacting a cell with a BZRP-R4 agonist. This method comprises the step of contacting a cell with an agonist of BZRP-R4. Thus, the subject invention provides methods of treating disorders including, but not limited to, schizophrenia, chronic stress, GAD, PD, GSP and PTSD. Other disorders which may be treated by agonists of BZRP-R4 25 include those associated with decreases in cell proliferation, e.g. developmental retardation. Furthermore, because BZRP-R4 is able to transport cholesterol into cells, BZRP-R4 agonists may also be used to increase cholesterol transport into cells. Diseases associated with cholesterol transport deficiencies include lipoidal adrenal hyperplasia, ovarian cysts, abnormal lipid deposits in steroidogenic cells. Disorders that reflect a requirement for cholesterol for myelin and myelination, 30 include Alzheimer's disease, multiple sclerosis, spinal cord injury, and brain development neuropathy. The methods of treating disorders associated with decreased levels of BZRP-R4 may be practiced by introducing agonists which stimulate the expression or the activity of BZRP-R4.

[0446] Additionally, disorders resulting from defective mitochondrial activity may be treated with an agonist to BZRP-R4. Defective mitochondrial activity may alternatively or additionally result in the generation of highly reactive free radicals that have the potential of damaging cells and tissues. These free radicals may include reactive oxygen species (ROS) such as superoxide, peroxynitrite and hydroxyl radicals, and other reactive species that may be toxic to cells and cause apoptosis. For example, oxygen free radical induced lipid peroxidation is a well-established pathogenic mechanism in central nervous system (CNS) injury such as that found in a number of degenerative diseases, and in ischemia (i.e., stroke). Diseases associated with altered mitochondrial function and apoptosis include: Alzheimer's Disease, diabetes mellitus, Parkinson's Disease, Huntington's disease, dystonia, Leber's hereditary optic neuropathy, schizophrenia, mitochondrial encephalopathy, lactic acidosis, and stroke.

[0447] A further preferred embodiment includes a method of inhibiting apoptosis of cells in culture. This method comprises the step of contacting a cell in culture with an agonist to BZRP-R4. Such methods are useful for culturing cells that are notoriously undergo apoptosis, such as primary neurons and lymphocytes.

[0448] In one embodiment, the level of BZRP-R4 in a cell may be increased by introducing nucleic acids encoding a BZRP-R4 polypeptide or biologically active fragment thereof into a targeted cell type. Vectors useful in such methods are known to those skilled in the art, as are methods of introducing such nucleic acids into target tissues.

[0449] Antibodies or other polypeptides capable of reducing or inhibiting the activity of BZRP-R4 may be provided as in isolated and substantially purified form. Alternatively, antibodies or other polypeptides capable of inhibiting or reducing the activity of BZRP-R4 may be recombinantly expressed in the target cell to provide a modulating effect. In addition, compounds which inhibit or reduce the activity of BZRP-R4 may be incorporated into biodegradable polymers being implanted in the vicinity of where drug delivery is desired. For example, biodegradable polymers may be implanted at the site of a tumor or, alternatively, biodegradable polymers containing antagonists/agonists may be implanted to slowly release the compounds systemically. Biodegradable polymers, and their use, are known to those of skill in the art (see, for example, Brem et al. (1991) J. Neurosurg. 74:441-446, which disclosure is hereby incorporated by reference in its entirety).

[0450] In another embodiment, the invention provides methods and compositions for detecting the level of expression of the mRNA encoding the protein of the invention. Quantification of mRNA levels of BZRP-R4 may be useful for the diagnosis or prognosis of diseases associated with an altered expression of the protein of the invention. Assays for the detection and quantification of the

mRNA encoding BZRP-R4 are well known in the art (see, for example, Maniatis, Fitzsch and Sambrook, Molecular Cloning; A Laboratory Manual (1982), or Current Protocols in Molecular Biology, Ausubel, F.M. et al. (Eds), Wiley & Sons, Inc., disclosures of which are hereby incorporated by reference in their entireties).

5 [0451] Polynucleotides probes or primers for the detection of BZRP-R4 mRNA can be designed from the cDNA of SEQ ID NO:19. Methods for designing probes and primers are known in the art. In another embodiment, the subject invention provides diagnostic kits for the detection of the mRNA of the protein of the invention in cells. The kit comprises a package having one or more containers of oligonucleotide primers for detection of the polynucleotides of the invention in PCR
10 assays or one or more containers of polynucleotide probes for the detection of the BZRP-R4 mRNA by in situ hybridization or Northern analysis. Kits may, optionally, include containers of various reagents used in various hybridization assays. The kit may also, optionally, contain one or more of the following items: polymerization enzymes, buffers, instructions, controls, or detection labels. Kits may also, optionally, include containers of reagents mixed together in suitable proportions for performing
15 the hybridization assay methods in accordance with the invention. Reagent containers preferably contain reagents in unit quantities that obviate measuring steps when performing the subject methods.

[0452] In another embodiment, the invention relates to methods and compositions for detecting and quantifying the level of the protein of the invention present in a particular biological sample. These methods are useful for the diagnosis or prognosis of diseases associated with altered
20 levels of the protein of the invention. Diagnostic assays to detect the protein of the invention may comprise a biopsy, in situ assay of cells from organ or tissue sections, or an aspirate of cells from a tumor or normal tissue. In addition, assays may be conducted upon cellular extracts from organs, tissues, cells, urine, or serum or blood or any other body fluid or extract.

[0453] Assays for the quantification of BZRP-R4 polypeptides may be performed according
25 to methods well known in the art. Typically, these assays comprise the steps of: contacting the sample with a ligand of the protein of the invention or an antibody (polyclonal or monoclonal) that specifically recognizes the protein of the invention or a fragment thereof and detecting the complex formed between the protein of the invention present in the sample and the ligand or antibody. Fragments of the ligands and antibodies may also be used in the binding assays, provided these fragments are
30 capable of specifically interacting with BZRP-R4 polypeptides. Further, ligands and antibodies which bind to BZRP-R4 may be labeled according to methods known in the art. Labels which are useful in the subject invention include, but are not limited to, enzymes labels, radioisotopic labels, paramagnetic labels, and chemiluminescent labels. Typical techniques are described by Kennedy, J. H., et al. (1976)

Clin. Chim. Acta 70:1-31; and Schurs, A. H. et al. (1977) Clin. Chim. Acta 81: 1-40, disclosures of which are hereby incorporated by reference in their entireties.

[0454] The subject invention also provides methods and compositions for the identification of metastatic tumor masses. In this aspect of the invention, the polypeptide or antibody that specifically binds a BZRP-R4 polypeptide or fragment thereof may be used as a marker for the identification of the metastatic tumor mass. Metastatic tumors which originated from the breast or liver may overexpress BZRP-R4 polypeptides, whereas newly forming tumors, or those originating from other tissues are not expected to bear BZRP-R4.

Protein of SEQ ID NO:22 (Internal designation Clone 420594 145-19-4-0-E7-F)

10 [0455] The cDNA of Clone 420594 145-19-4-0-E7-F (SEQ ID NO:21) encodes the 106 amino acid Scolakin protein of SEQ ID NO:22 comprising the amino acid sequence:

MPFLDIQKRFGLNIDRWLTIQSCEQPYKMAGRCHAFKEWIECAHGIGYTRAECKIEYDDF
VECLLRQKTMRRAGTIRKQRDKL¹⁵IKEGKYTPPPHHIGKGEWP. Accordingly, it will be appreciated that all characteristics and uses of polypeptides of SEQ ID NO:22 described throughout the present application also pertain to the polypeptides encoded by the nucleic acids comprising the human cDNA in Clone 420594 145-19-4-0-E7-F. In addition, it will be appreciated that all characteristics and uses of the polynucleotides of SEQ ID NO:21 described throughout the present application also pertain to the nucleic acids included in Clone 420594 145-19-4-0-E7-F. Also preferred are fragments having a biological activity as described therein and the polynucleotides 20 encoding the fragments.

[0456] Scolakin, the protein of SEQ ID NO:22, represents a novel variant form of the 15 kilodalton NADH dehydrogenase subunit NDUFS5 (Genbank entry AF020352).

[0457] NADH dehydrogenase (NADH:ubiquinone oxidoreductase, complex I) is the first multisubunit inner membrane protein complex in a chain of three complexes that make up the 25 mitochondrial electron transport chain. The mitochondrial electron transport chain is responsible for the transport of electrons from NADH to oxygen and the coupling of this oxidation to the synthesis of ATP (oxidative phosphorylation) which provides the energy source for driving a cell's many energy-requiring reactions. NADH dehydrogenase accomplishes the first step in this process by accepting two electrons from NADH and passing them through a flavin molecule to ubiquinone, which transfers the 30 electrons to the second enzyme complex in the chain. It contains a number of prosthetic groups which are involved in the transfer of electrons, namely, flavin mononucleotide (FMN) and at least six iron-sulphur clusters (binuclear and tetranuclear). NADH dehydrogenase is the largest of the three

complexes with an estimated mass of 800 kDa comprising about 41 polypeptides of widely varying size and composition. Seven of these complex I polypeptides are encoded by mitochondrial DNA while the remaining are nuclear gene products that are imported into the mitochondria. The polypeptide composition of NADH dehydrogenase is very similar in a variety of mammalian species 5 including rat, rabbit, cow, and man. In man, deficiency of NADH dehydrogenase is one of the most frequent causes of human encephalomyopathies. Defects and altered expression of this enzyme have also been reported to be associated with other disease conditions, i.e., neurodegenerative diseases and cancer. In addition, NADH dehydrogenase reduction of the quinone moiety in chemotherapeutic agents such as doxorubicin is believed to contribute to the antitumor activity and/or mutagenicity of 10 these drugs. Defects in mitochondrial-encoded subunits of complex I have been described. However a relatively small percentage of human complex I deficiency is associated with mitochondrial DNA mutations which suggests that most of the corresponding mutations affect expression or activity of nuclear-encoded subunits of complex I.

[0458] Scolakin, the protein of the invention, belongs to the category of the nuclear-encoded 15 subunits of complex I. This subunit is part of the iron-sulphur protein fraction (IP) during complex I purification and it is believed to participate to the redox reactions catalysed by the iron-sulphur centers of this complex. NDUFS5 contains four cysteine residues that form one of the iron-sulphur centers that function in electron transport. Interestingly, Scolakin contains a fifth cysteine due to a glycine to cysteine substitution at position 23 on SEQ ID NO:22. Scolakin is widely expressed with a relatively 20 higher expression in heart, brain, skeletal muscle, liver, kidney and fetal heart and brain, all tissues with high metabolic activities, which are often clinically affected in complex I-deficient patients. In addition, Scolakin expression is high in cancerous tissues and immortalized cells lines, which also have high metabolic requirements. Scolakin plays a role in myopathies, neurodegenerative diseases and cancer.

25 [0459] An embodiment of the invention is directed to a composition comprising a Scolakin polypeptide sequence of SEQ ID NO:22.

[0460] A further embodiment of the invention is directed to a composition comprising a Scolakin polypeptide fragment having biological activity.

30 [0461] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence of SEQ ID NO:21 encoding a Scolakin polypeptide.

[0462] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence encoding a Scolakin polypeptide fragment having biological activity.

[0463] A further embodiment of the invention is directed to a method of screening test substances for modulators of Scolakin expression comprising the steps of: i) contacting a cell with a substance to be tested; and ii) comparing Scolakin expression in the cell after exposure to the test substance to that of an untreated control cell.

5 [0464] In one embodiment, a sequence encoding SEQ ID NO:22 bearing G to T and C to T substitutions at nucleotide positions 152 and 398 of SEQ ID NO:21 corresponding to positions 23 and 105, and resulting in the substitution of a glycine residue by a cysteine at position 23 and an arginine residue by a tryptophane at position 105, respectively, can be used for DNA genotyping. Genotyping this locus could be of interest, e.g., in DNA fingerprinting for paternity studies or forensic analyses. It
10 could also be used for genetic association studies, preferably in pathologies relating to mitochondrial disorders, especially those with isolated complex I deficiency.

[0465] In another embodiment, the polynucleotide sequence of the invention can be used in pharmacogenomic applications in order to aid in the choice of the ideal drug (e.g. an agonist or an antagonist of Scolakin), or dosage of a drug, for the treatment of a condition or disease in a patient.
15 For example, in one embodiment, the invention provides a method of genotyping a patient to determine the identity of the nucleotides encoding the amino acids at positions 23 and 105 of Scolakin, and administering to the patient a drug or a dosage of the drug that has been established to be preferentially efficacious in those with cysteine and tryptophane residues at positions 23 and 105, respectively (e.g. because of preferential binding of the drug to the isoform of the protein with cysteine
20 and tryptophane at these positions). In another embodiment, the patient is genotyped for the nucleotides encoding amino acid positions 23 and 105, a drug is not administered, e.g. because side effects are known to be associated with the administration of the drug to individuals with cysteine and tryptophane at positions 23 and 105, respectively.

[0466] A preferred aspect of the invention is a host cell recombinant for polynucleotides
25 encoding a Scolakin polypeptide or a biologically active fragment thereof. An additional preferred aspect is a host cell recombinant for polynucleotides capable of directing Scolakin expression. Preferably, the polynucleotides capable of directing Scolakin expression are located in the 5' regulatory region of the Scolakin gene. Further preferably, these polynucleotides are located within
30 500 base pairs of the Scolakin coding region. These polynucleotides preferably comprise a promoter sequence. Techniques known in the art for introducing polynucleotide sequences to endogenous sequences are described in U.S. Patent 5641670 and PCT WO9629411, which disclosures are hereby incorporated by reference in their entireties. Scolkin protein produced by said host cell may be used for in vitro detection and purification methods as well as diagnosis and in vivo applications.

[0467] Another embodiment relates to methods of producing Scolakin polypeptides. The protein of the invention can be produced in host cells that have been transfected with an appropriate expression vector comprising a nucleic acid sequence coding for the protein of the invention. Introduction into a host cell of such expression vector for Scolakin can be performed in a variety of ways, including but not limited to calcium or lithium chloride treatment, electroporation, or lipofection. Any of a wide variety of expression systems can be used to provide the recombinant proteins. Suitable expression vehicles include, but are not limited to plasmids, viral particles or baculovirus for insect cells. The expression vehicle can be integrated into the host cell genome. Optionally, an inducible expression vector can be used to achieve tight controlled expression of the gene in the host cell. The recombinant protein can be recovered from the host cell and purified by any technique well known to those skilled in the art. Preferably, an antibody directed against the protein of the invention or part thereof can be bound to a chromatographic support to form an affinity chromatography column. Alternatively, Scolakin can be chemically synthesized using solid-phase techniques.

[0468] In a further embodiment, the invention provides compositions and methods using Scolakin or fragments thereof to label mitochondria in order to visualize any change in number, topology or morphology of this organelle, for example in association with a mitochondria-related human disorder, such as mitochondrial encephalomyopathies, Alpers disease, hereditary spastic paraplegia, neuroleptic malignant syndrome (NMS) or the Rett syndrome. For example, the protein may be rendered easily detectable by inserting the cDNA encoding the protein of the invention into a eukaryotic expression vector in frame with a sequence encoding a tag sequence. A preferred embodiment of the invention is a method to detect mitochondria comprising the step of contacting a monoclonal or polyclonal antibody specific for Scolakin with a cell. Preferably, the cell is fixed or otherwise permeable to antibodies. Preferably, the antibody is primarily or secondarily linked to a fluorescent, radioactive, or otherwise detectable compound.

[0469] In another embodiment, the polynucleotide of SEQ ID NO:21 or fragments thereof may be used to target recombinant DNA molecules to the mitochondria. For instance, a chimeric polynucleotide composed of the polynucleotide sequence of the invention recombinantly or chemically fused to a polynucleotide of therapeutic interest would allow the delivery of the therapeutic polynucleotide specifically to the mitochondria. Such chimeric molecules would be of particular interest in gene therapy to restore or modulate mitochondrial activities for the treatment and/or the prevention of disorders due to mitochondrial dysfunction, including, but not limited to, those mentioned above.

[0470] In another embodiment, the invention provides compositions and methods for detecting the level of expression of the mRNA encoding the protein of the invention in a mammal, preferably a human. Quantification of mRNA levels of Scolakin may be useful for the diagnosis of diseases or conditions correlated with abnormal expression of the protein, or to monitor regulation of Scolakin during therapeutic intervention as described herein. Assays for the detection and quantification of mRNA are well known in the art. Preferred method comprises the steps of: isolating RNA from a biological sample of a subject, measuring Scolakin mRNA level by quantitative RT-PCR, and comparing the expression in the subject sample to that of a control sample. Polynucleotide probes or primers used for the detection of Scolakin mRNA by hybridization or PCR amplification may be designed from the polynucleotide of SEQ ID NO:21 by methods well known in the art.

[0471] In another embodiment, the invention relates to methods and compositions for detecting Scolakin and quantifying its level of expression in a biological sample. These methods may be useful for the diagnosis of conditions or diseases characterized by altered or abnormal expression of the protein of the invention, or in assays to monitor subjects being treated with Scolakin, agonists or antagonists. A preferred method comprises contacting an antibody which specifically binds to Scolakin with a biological sample from a mammalian subject, preferably human, and determining the level of Scolakin in the subject sample compared to a control level representative of a healthy subject, wherein an altered or abnormal level of Scolakin in the subject sample relative to the control level indicates that the subject has the disease or is at an elevated risk of developing the disease. The antibody used can be either monoclonal or polyclonal and can be labeled directly or indirectly for quantification of immune complexes by methods well known to those skilled in the art, for example by ELISA or radioimmunoassays. Diagnostic assays to detect the protein of the invention may comprise a biopsy, in situ assay of cells from organ or tissue sections, or an aspirate of cells from a tumor or normal tissue. In addition, assays may be conducted upon cellular extracts from organs, tissues, cells, urine, or serum or blood or any other body fluid or extract.

[0472] Another embodiment of the invention is directed to a method to treat or prevent a defect in mitochondrial complex I of a mammal, preferably a human, comprising administering to said mammal the protein of the invention or fragments thereof or polynucleotides encoding the present protein. Expression vectors derived from retroviruses, adenovirus, herpes, or vaccinia viruses may be used for the delivery of the polynucleotide sequences of the invention to the targeted organ, tissue or cell populations. Many methods for introducing vectors into cells or tissues are available and equally suitable for use either in vivo or ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the subject and clonally propagated for autologous transplant back into the same

subject. Delivery by transfection and liposome injections may be achieved using methods which are well known in the art. Pharmaceutical compositions comprising Scolakin polypeptides can be administered alone or in combination with at least one other agent, such as stabilizing compound, which may be administered in any sterile biocompatible carrier. They may be administered by any 5 number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intranasal, enteral or rectal means. The compositions may be administered to the subject alone or in combination with other agents, drugs or hormones. Preferably, the diseases associated with defective complex I activity to be treated or prevented using the present method include, but are not limited to, fatal neonatal lactic 10 acidosis, myopathy with exercise intolerance and lactic acidosis, MELAS (mitochondrial myopathy, encephalopathy, lactic acidosis, and stroke-like episodes), MERFF (myoclonus epilepsy, and ragged red fibres), encephalomyopathy of childhood and adult life, Leigh syndrome, Alpers and Parkinson's diseases.

[0473] An embodiment of the invention provides for a method of screening test substances for 15 modulators of Scolakin expression. This method comprises the steps of: i) contacting a cell with a test substance; and ii) comparing Scolakin expression in the cell after exposure to the test substance to that of an unexposed control cell. Scolakin expression is determined by methods common to the art or included herein, by detecting Scolakin polynucleotides or polypeptides. An example of this method comprises the steps of: i) culturing two equivalent cell samples; ii) adding a test substance to one of 20 the cultures and not the other; iii) harvesting both cultures at a specified time; iv) purifying the mRNA from each sample of cells; v) comparing the level of Scolakin mRNA in each sample by Northern blot, RT-PCR, or another method common to the art. The invention provides for design and use of specific polynucleotide probes and primers, as discussed herein. An additional example comprises the steps of: i) having two equivalent cultures of cells; ii) adding a test substance to one of the cultures and not the 25 other; iii) harvesting both cultures; iv) purifying the protein from each sample of cells; v) comparing the level of Scolakin polypeptides in each sample by enzyme-linked immunoabsorbent assay (ELISA), western blot, radioimmunoassay (RIA), or another method common to the art. The invention provides for design and use of specific antibodies and antibody fragments, as discussed herein. Substances that increase Scolakin expression (agonists or activators) may be used to increase complexe I activity. 30 Substances that decrease Scolakin expression (antagonists or inhibitors) may be used to treat or prevent cancers. Methods utilizing Scolakin agonists and antagonists are included herein.

[0474] A preferred embodiment of the invention provides a method of screening for test substances that bind Scolakin polypeptides. This method comprises the steps of: i) contacting a test

substance with a Scolakin polypeptide or fragment thereof under conditions that allow binding; and ii) detecting the binding of the test substance by methods common to the art (e.g., competitive antibody-based methods such as coimmunoprecipitation and Western blotting). Included in this method are test substances that are conjugated to an antibody, antibody fragment, cell-type specific ligand or a portion 5 thereof.

[0474] A further preferred embodiment of the invention provides a method of screening test substances that bind to Scolakin for agonists of Scolakin activity, comprising: i) contacting a cell with the substance to be tested; and ii) comparing Scolakin biological activity after exposure to the test substance to that of an unexposed control cell. Measure of Scolakin biological activity may be 10 assessed indirectly by measuring the enzymatic activity of complex I as described by Loeffen et al. (1999, J. Inher. Metab. Dis. Vol. 22, pp 19-28 which disclosure is hereby incorporated by reference in its entirety), an increased in complex I activity in the test sample compared to the control sample being indicative of an activating effect of the test substance on Scolakin.

[0475] A further preferred embodiment of the invention provides a method of screening test 15 substances that bind to Scolakin for antagonists of Scolakin activity. This method comprises the steps of: i) contacting a cell with a test substance; and ii) comparing Scolakin biological activity after exposure to the test substance to that of an unexposed control cell. Detection of Scolakin biological activity may be detected by measuring complex I activity as mentioned above, a decrease in complex I activity in the test sample compared to the control sample being indicative of an inhibitory effect of the 20 test substance on Scolakin activity.

[0476] It is yet another object of the invention to provide methods of treating or preventing cancers associated with increased complex I activity. This can be achieved by decreasing Scolakin expression or activity in vivo using antibodies, ribozymes or antisense vectors or oligonucleotides. Alternatively, an antagonist of Scolakin activity or expression isolated as described above can be used, 25 such method comprising the step of administering to a subject an antagonist of Scolakin expression or activity. Scolakin antagonists in a physiologically acceptable solution may be delivered by methods common to the art, such as orally or parenterally. Preferably, the Scolakin antagonist is delivered to a specific cell type, for example, by conjugating the antagonist to a cell-type specific targeting moiety (e.g., a ligand or antibody fragment). This method is useful for prevention and treatment of cancers 30 including, but not limited to, cancers of the heart, ovaries, colon, kidney, bladder, prostate, pancreas, brain, stomach, breast, lung, liver and leukemias.

Protein of SEQ ID NO:24 (Internal designation Clone 119658 105-067-2-0-H4-F_1)

[0477] The cDNA of Clone 119658 105-067-2-0-H4-F_1 (SEQ ID NO:23) encodes the Docking Of Vesicles (DOV) protein of SEQ ID NO:24, comprising the amino acid sequence: MASRSSDKDGDSVHTASEVPLTPRTNSPDGRRSSSDTSKSTYSLTRRISSLESRRPSSPLIDIKPIE FGVLSAKKEPIQPSVLRRTYNPDDYFRKFEPHLYSLDSNSDDVDSLTDDEEILSKYQLGMQHFST 5 QYDLLLHNHLTVRVIEARDLPPPISHDGSRQDMAHSNPYVKICLLPDQKNSKQTGVKRKTQKP VFEERYTTEIPFLEAQRRRTLLLTVVDFDKFSRHCVIGKVSVPLCEVDLVKGHHWWKALIPSSQ NEVELGELLLSBNYLPSAGRLNVDVIRAKQLLQTDVSQGSDPFVKIQLVHGLKLVTKKTSFL RGTIDPFYNESFSFKVPQEELENASLVFTVFGHNMKSSNDFIGRIVIGQYSSGPSETNHWRML NTHRTAVEQWHSLRSRAECDRVSPASLEVT. Accordingly, it will be appreciated that all 10 characteristics and uses of the polypeptides of SEQ ID NO:24 described throughout the present application also pertain to the polypeptides encoded by the nucleic acids comprising the human cDNA in Clone 119658 105-067-2-0-H4-F_1. In addition, it will be appreciated that all characteristics and uses of the polynucleotides of SEQ ID NO:23 described throughout the present application also pertain to the nucleic acids included in Clone 119658 105-067-2-0-H4-F_1. A preferred embodiment of the 15 invention is directed toward the compositions of SEQ ID NO:23, SEQ ID NO:24, and Clone 119658 105-067-2-0-H4-F_1. Also preferred are polypeptide fragments having a biological activity as described herein and the polynucleotides encoding the fragments.

[0478] The sequence of DOV contains similarities to proteins of the Synaptotagmin family. Like Synaptotagmin, DOV is a vesicular transmembrane protein that participates in vesicle docking 20 and fusion. A calcium-dependent conformational change allows the cytoplasmic portion of DOV to dock with specific proteins on the plasma membrane. Vesicle docking results in release of vesicle contents to the extracellular space. DOV polypeptides are required for neurotransmitter release from a presynaptic neuron in response to an excitatory calcium flux in that cell.

[0479] A preferred embodiment of the invention includes a composition comprising a DOV 25 polypeptide sequence of SEQ ID NO:24.

[0480] A preferred embodiment of the invention includes a composition comprising a DOV polypeptide fragment having biological activity.

[0481] A preferred embodiment of the invention includes a composition comprising a polynucleotide sequence of SEQ ID NO:23 encoding a DOV polypeptide.

30 [0482] A preferred embodiment of the invention includes a composition comprising a polynucleotide sequence encoding a DOV polypeptide fragment having biological activity.

[0483] A preferred embodiment of the invention includes an excitation-secretion uncoupling peptide (ESUP). Preferred DOV ESUPs include the peptide sequences: LLHNHLTDRVIEARDLPPPISHGSRQDMAHSNPYVKICLLPDQKNSKQTGVKRKTQKPVFEE RYTFEIPFLEAQRTLLLTVVDFDKFSRHCVIGKVS; 5 GRLNVDVIRAKQLLQTDVSQGSDFVKIQLVHGLLVTKKTSFLRGTIIDPFYNESFSFKVPQE ELENASLVFTVFGHNMKSSNDFIGRIVIG; and effective ESUP fragments of said peptides.

[0484] A preferred embodiment of the invention provides a method of screening for inhibitors of DOV polypeptides comprising the steps of: contacting a test substance with a cell and detecting DOV biological activity in the exposed cell to that in an unexposed control cell. Preferred cells are 10 neurons. Preferred antagonists include DOV-specific antibodies and fragments thereof, ESUPs, and small molecules.

[0485] A preferred embodiment of the invention provides a method of using inhibitors of DOV polypeptides as anticonvulsants and to prevent pain. These methods comprise the step of administering an effective amount of a DOV inhibitor in a physiologically acceptable composition to 15 an individual in need of treatment.

[0486] A preferred embodiment provides a method to purify vesicles comprising the steps of: i) contacting an antibody specific for the cytoplasmic portion of DOV polypeptide with a biological sample under conditions that allow antibody binding and ii) removing contaminating materials not bound to the antibody. Preferably, such antibody is primarily or secondarily attached to an insoluble 20 matrix to enable purification.

[0487] A preferred embodiment provides a method to detect vesicles comprising the step of contacting an antibody specific for DOV polypeptide with a biological sample under conditions that allow binding and ii) detecting the antibody. Preferred antibodies are covalently labeled with fluorescent, radioactive, or otherwise detectable compounds.

25 [0488] A preferred embodiment provides a method to detect vesicle extrusion comprising the step of contacting an antibody specific for the vesicular portion of DOV with a cell. Preferred antibodies are covalently labeled with fluorescent, radioactive, or otherwise detectable compounds. Preferred cells are neurons.

[0489] A preferred embodiment of the invention provide a method to diagnose a neurological 30 disorder comprising the steps of: obtaining a biological test sample from an individual; detecting the level of DOV expression in the test sample; comparing to the level of DOV expression in the test sample to that of control sample(s). Neurological disorders indicated by this method include:

Paroxysmal Kinesigenic Choreoathetosis (PKC), paralysis (partial and complete), and mood disorders (depression, bipolar disorder, and schizophrenia).

[0490] A preferred embodiment is a method to restore vesicular docking and fusion abilities to a cell comprising the step of introducing DOV polypeptides to a cell. Preferably, this method is applied to treatment and prevention of PKC, partial and complete paralysis, and mood disorders such as depression, bipolar disorder, and schizophrenia. Preferably, DOV polypeptides are delivered by introducing a polynucleotide construct comprising an expression control element operably linked to a polynucleotide encoding a DOV polypeptide to a cell.

[0491] A preferred aspect of the invention is a host cell recombinant for polynucleotides encoding a DOV polypeptide or a biologically active fragment thereof. An additional preferred aspect is a host cell recombinant for polynucleotides capable of directing DOV expression. Preferably, the polynucleotides capable of directing DOV expression are located in the 5' regulatory region of the DOV gene. Further preferably, these polynucleotides are located within 500 base pairs of the DOV coding region. These polynucleotides preferably comprise a promoter sequence. Techniques known in the art for introducing polynucleotide sequences to endogenous sequences are described in U.S. Patent 5641670 and PCT WO9629411, which disclosures are hereby incorporated by reference in their entireties. DOV protein produced by said host cell may be used for in vitro detection and purification methods as well as diagnosis and in vivo applications.

[0492] In one aspect, the invention includes a DOV polypeptide composition for use in delivering a second composition, preferably nucleic acids, polypeptides, or small molecules, as therapeutic drugs, to target biological cells either in vitro or in vivo. The second composition may, if desirable, be covalently or non-covalently attached or fused to the DOV polypeptide. The DOV polypeptide composition may further comprise artificial lipids or viral components to facilitate delivery of the second molecule by liposomes or lipid vesicles. Methods for using DOV polypeptides in these methods are known in the art and include U.S. Patents 6,074,844 and 6,099,857, which disclosures are hereby incorporated by reference in their entireties. In a preferred embodiment, DOV polypeptides are used to facilitate delivery of a second composition, e.g., liposome-mediated DNA transfection, to cells in culture, preferably neuronal cells.

[0493] DOV polypeptides are also useful in methods of inhibiting the release of neurotransmitters by preventing the docking and/or fusing of a presynaptic vesicle to the presynaptic membrane. These polypeptides may be referred to as excitation-secretion uncoupling peptides (ESUPs). Fragments of DOV having this blocking activity can be identified using methods known in

the art (See e.g., U.S. Patents 6,090,631 and 6,169,074 incorporated by reference in their entireties). ESUPs of the present invention comprise synthetic and purified DOV peptide fragments which correspond in primary structure to peptides which serve as binding domains for the assembly of a ternary protein complex ("docking complex") which is critical to neuronal vesicle docking with the 5 cellular plasma membrane prior to neurotransmitter secretion. For optimal activity, ESUPs of the invention have a minimum length of about 20 amino acids and a maximal length of about 100 amino acids, although they may be larger or smaller. Preferred DOV ESUPs for use in inhibiting the release of neurotransmitters include those comprising the sequence:
LLHNHLTVRVIEARDLPPPIHDGSRQDMAHSNPYVKICLLPDQKNSKQTGVKRKTQKPVFEE
10 RYTFEIPFLEAQRTLLLTVVDFDKFSRHCVIGKVS;
GRLNVDVIRAKQLLQTDVSQGSDPFVKIQLVHGLKLVTKKTSFLRGTIIDPFYNESFSFKVPQE
ELENASLVFTVFGHNMKSSNDFIGRIVIG; and effective ESUP fragments of said peptides. DOV ESUPs are preferably covalently linked to a targeting moiety. Preferred targeting moieties comprise a ligand to a cell-surface binding site present on a specific cell type (e.g., neuron) that is capable of 15 functionally interacting with the binding site. Preferred targeting moieties are nerve growth factor and functional derivatives thereof. Alternatively, DOV ESUPs may be expressed in a cell by introducing polynucleotides encoding DOV ESUPs to the cell. Methods of introducing polynucleotides to a cell are known to those skilled in the art, as discussed herein and in U.S. Patents 6117454 and 6180602, which disclosures are hereby incorporated by reference in their entireties.

20 [0494] Another embodiment of the invention provides a method of screening for antagonists of DOV polypeptides. This method comprises the steps of: contacting a test substance with a cell and detecting DOV biological activity in the exposed cell to that in an unexposed control cell. Preferred cells are neurons. DOV polypeptides are required for efficient neurotransmitter release. Thus, DOV activity can be determined by detecting the level of neurotransmitter released from exposed cells 25 compared to that from unexposed cells. Such assays are common to the art and include antibody-based assays (e.g., ELISA), binding competition assays (e.g., to a specific neurotransmitter receptor), and colormetric assays. Preferred antagonists include DOV-specific antibodies and fragments thereof, ESUPs, and small molecules.

[0495] DOV antagonists may be used to inhibit or treat pain. Peptide antagonists in particular 30 may be used according to U.S. Patent 5,989,545 (disclosure of which is hereby incorporated by reference in its entirety) by substituting the polypeptides of the present invention for neurotoxin. In addition, DOV antagonists of the invention may be used as anticonvulsants. Anticonvulsants effectively reduce infarct size in cases of stroke and are effective therapy for seizures, such as those

resulting from epilepsy or exposure to neurotoxins. These methods comprise the step of administering an effective amount of DOV antagonist in a physiologically acceptable composition to an individual in need of treatment.

[0496] DOV is a component of vesicles, thus, antibodies to DOV are useful in the detection of 5 vesicles, preferably neuronal vesicles transporting neurotransmitters. DOV-specific antibodies can be used during purification of vesicles. This method comprises the steps of: i) contacting an antibody specific for the cytoplasmic portion of DOV polypeptide with a biological sample under conditions that allow antibody binding and ii) removing contaminating materials not bound to the antibody. Preferably, such antibody is primarily or secondarily attached to an insoluble matrix to enable 10 purification. DOV-specific antibodies may also be used as a marker for vesicles. For example, vesicles can be detected in assays such as immunohistochemistry. The vesicular portion of DOV appears on the surface of the cell following extrusion of vesicle contents at the plasma membrane. Thus, antibodies specific for the vesicular portion of DOV are useful for the detection and monitoring 15 of vesicle extrusion. This method comprises the step of contacting an antibody specific for the vesicular portion of DOV with a cell. Preferred antibodies are detectably labeled, for example, with a fluorescent, luminescent, or radiolabeled compound.

[0497] Detection of DOV expression (mRNA or protein) levels or mutated forms of DOV is further useful in the determination or diagnosing of whether someone is at risk of developing or has a neurological disorder, such as Paroxysmal Kinesigenic Choreaathetosis (PKC), paralysis (partial or 20 complete), or mood disorders (e.g., depression, bipolar disorder, schizophrenia). This method comprises the steps of: obtaining a biological test sample from an individual; detecting the level of DOV expression in the test sample; comparing to the level of DOV expression in the test sample to that of control sample(s). A decreased level in expression of DOV, mRNA or protein, as compared to an individual without a neurological disorder indicates the individual has or is at risk of developing the 25 disorder.

[0498] DOV polypeptides may further be introduced to a cell that is defective in vesicle docking and fusion. This method comprises the step of introducing DOV polypeptides to a cell such that vesicle docking and fusion are allowed. DOV polypeptides may be introduced to a cell by introducing a polynucleotide construct comprising an expression control element operably linked to 30 the coding sequence of DOV to a cell by methods discussed herein. Preferred cells are neurons. DOV polypeptides are preferably applied to treatment of PKC, paralysis (partial or complete), or mood disorders (e.g., depression, bipolar disorder, schizophrenia).

Protein of SEQ ID NO:26 (Internal designation Clone 500706437_204-5-2-0-C9-F)

[0499] The cDNA of Clone 500706437_204-5-2-0-C9-F (SEQ ID NO:25) encodes placentalin of SEQ ID NO:26, comprising the amino acid sequence: MDPARPLGLSILLFLTEAALGDAAQEPTGNNAEICLLPLDYGPCRALLRYYYDRYTQSCRQ 5 FLYGGCEGNANNFYTWEACDDLAGG. Accordingly, it will be appreciated that all characteristics and uses of the polypeptides of SEQ ID NO:26 described throughout the present application also pertain to the polypeptides encoded by the nucleic acids comprising the human cDNA in Clone 500706437_204-5-2-0-C9-F. In addition, it will be appreciated that all characteristics and uses of the polynucleotides of SEQ ID NO:25 described throughout the present application also pertain to the 10 nucleic acids included in Clone 500706437_204-5-2-0-C9-F. A preferred embodiment of the invention is directed toward the compositions of SEQ ID NO:26, SEQ ID NO:25, and Clone 500706437_204-5-2-0-C9-F. Preferred placentalin polypeptides for uses in the methods described below include the polypeptides comprising the amino sequence of: 15 AQEPTGNNAEICLLPLDYGPCRALLRYYYDRYTQSCRQFLYGGCEGNANNFYTWEACDDL AGG. Also preferred are polypeptide fragments having a biological activity as described herein and the polynucleotides encoding the fragments.

[0500] The protein of the present invention, placentalin, is a variant of the Tissue Factor Pathway Inhibitor 2 (TFPI-2, GenBank accession number P48307). The 84 amino-terminal amino-acids are identical between TFPI-2 and placentalin, and the 4 carboxyl-terminal amino-acids are 20 unique to placentalin. Placentalin displays one signal peptide (MDPARPLGLSILLFLTEAALGDA) and one Kunitz inhibitor domain (CLLPLDYGPCRALLRYYYDRYTQSCRQFLYGGCEGNANNFYTWEACDDL). Placentalin lacks the main heparin-binding site, located at the carboxyl-terminal extremity of TFPI-2. A wide 25 variety of cell types, including endothelial cells, keratinocytes and fibroblasts, synthetize and secrete placentalin under normal conditions. Moreover, placentalin is highly expressed in umbilical vein endothelial cells and its expression level negatively correlates with malignancy-grade of tumors.

[0501] Placentalin is a broad-spectrum serine protease inhibitor that inhibits plasmin, trypsin, plasma kallikrein, factor XIa, tissue factor-factor VIIa complex, chymotrypsin, factor IXa-polylysine and cathepsin G. Placentalin plays a major role both in regulation of extracellular matrix (ECM) 30 degradation and in regulation of the blood coagulation pathway. In peculiar, placentalin strongly inhibits the tissue factor-factor VIIa complex, the cellular initiator of the extrinsic coagulation pathway, and plasmin, thereby inhibiting processing of pro-MMPs (matrix metalloproteases)

responsible of ECM degradation. Unlike TFPI-2, placentalin is a circulating variant that is not bound to the heparin molecules of the ECM.

[0502] An embodiment of the invention is directed to a composition comprising a placentalin polypeptide sequence of SEQ ID NO:26.

5 [0503] A further embodiment of the invention is directed to a composition comprising a placentalin polypeptide fragment having biological activity of binding or inhibiting a serine protease.

[0504] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence of SEQ ID NO:25 encoding a placentalin polypeptide.

10 [0505] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence encoding a placentalin polypeptide fragment having biological activity.

[0506] A further embodiment of the invention is directed to a composition comprising an antibody recognizing a placentalin polypeptide sequence of SEQ ID NO:26 or a placentalin polypeptide fragment having biological activity. Preferably, the antibody recognizes an epitope comprising one or more of the 4 C-terminal amino-acids of placentalin, and binds to placentalin but 15 not to TFPI-2.

[0507] A preferred aspect of the invention is a host cell recombinant for polynucleotides encoding a placentalin polypeptide or a biologically active fragment thereof. An additional preferred aspect is a host cell recombinant for polynucleotides capable of directing placentalin expression. Preferably, the polynucleotides capable of directing placentalin expression are located in the 5' 20 regulatory region of the placentalin gene. Further preferably, these polynucleotides are located within 500 base pairs of the placentalin coding region. These polynucleotides preferably comprise a promoter sequence. Techniques known in the art for introducing polynucleotide sequences to endogenous sequences are described in U.S. Patent 5641670 and PCT WO9629411, which disclosures are hereby incorporated by reference in their entireties. Placentalin protein produced by said host cell may be 25 used for in vitro detection and purification methods as well as diagnosis and in vivo applications.

[0508] Another embodiment relates to a method of producing placentalin polypeptides comprising the steps of: i) transfecting a mammalian host cell with a recombinant expression vector comprising a polynucleotide of the present invention, and ii) purifying the produced protein. The purification of the protein can be done following any technique well-known to those skilled in the art. 30 Preferably, an antibody directed against placentalin or part thereof may be bound to a chromatographic

support to form an affinity chromatography column. Even more preferably, the antibody binds to placentalin but not to TFPI-2.

[0509] A preferred embodiment of the invention is a method of using placentalin to bind serine proteases. This method comprises the step of contacting a placentalin polypeptide or active 5 fragment thereof with a serine protease under conditions that allow placentalin binding, whereby binding inhibits the activity of said serine protease. Preferred serine proteases include plasmin, trypsin, plasma kallikrein, factor XIa, tissue factor-factor VIIa complex, chymotrypsin, factor IXa-polylysine or cathepsin G. This method may be applied to *in vitro* uses (e.g., protease purification and detection and prevention of protein degradation) as well as *in vivo* methods (e.g., inhibition of target 10 protease-dependent pathologies such as blood clots and metastasis).

[0510] A preferred embodiment of the invention is a method of purifying serine proteases. This method comprises the steps of: contacting a placentalin polypeptide or protease-binding fragment thereof with a serine protease under conditions that allow binding; removing contaminants; and eluting the serine protease with more stringent conditions. Preferably, the placentalin peptide is immobilized 15 on a solid or semi-solid matrix to facilitate washing of sample to remove contaminants. Preferred serine proteases include plasmin, trypsin, plasma kallikrein, factor XIa, tissue factor-factor VIIa complex, chymotrypsin, factor IXa-polylysine or cathepsin G. These may be purified from common biological fluids such as cell culture media and body fluids (e.g., serum, blood, lymph, or cerebrospinal fluid). Purified proteases are useful in biological assays and techniques such as removal 20 of adherent cells from a culture dish and site-directed peptide design.

[0511] A further embodiment of the present invention relates to a method of using placentalin or a protease-binding fragment thereof to remove contaminating proteases from a sample comprising the steps of: i) binding placentalin or part thereof to a chromatographic support, either alone or in combination with other protease inhibitors, and ii) running through the column the sample containing 25 the undesirable protease. The sample may be, e.g., a protein preparation or a sample prepared from bodily fluid. Preferably, the sample is prepared from blood.

[0512] An embodiment of the present invention relates to a method of using placentalin or part thereof to inhibit contaminating proteases in a sample. Such method comprises the step of adding a protease-inhibiting amount of placentalin polypeptide to a solution under conditions that allow 30 placentalin activity. Preferably, the placentalin polypeptide is comprised in a “cocktail” of protease inhibitors that is able to inhibit a wide range of proteases without knowing the specificity of any of the

proteases. Such protease inhibitor cocktails are widely used in laboratory assays to prevent degradation of any protein sample by contaminating proteases.

[0513] A preferred embodiment of the inventions is a method of detecting serine proteases. This method comprises the step of contacting a placentalin polypeptide or active fragment thereof with 5 a serine protease and detecting the presence of said serine protease by detecting placentalin. Preferably, the placentalin polypeptide is detectably labeled with, for example, a fluorescent, luminescent, or radioactive compound. Preferred serine proteases to be detected according to this method include plasmin, trypsin, plasma kallikrein, factor XIa, tissue factor-factor VIIa complex, chymotrypsin, factor IXa-polylysine or cathepsin G. Preferably, serine proteases are detected in 10 common biological fluids such as cell culture media and body fluids. This method may be applied to quantifying the level of serine protease expression in a cell sample or individual. This information may be useful in determination of tumor malignancy or diagnosis of blood coagulation disorders.

[0514] An additional preferred embodiment of the invention is a method of binding placentalin polypeptides with a placentalin-specific antibody or placentalin-binding fragment thereof. 15 This method comprises the step of contacting a placentalin polypeptide with a placentalin-binding antibody or placentalin-binding fragment thereof under conditions that allow binding. This method may be applied to detection and purification of placentalin, as well as modifying placentalin function. These aspects are discussed in detail herein.

[0515] Another embodiment of the present invention relates to a method of diagnosing the 20 malignancy of a tumor comprising the steps of: i) contacting a biological sample with a compound that specifically binds to placentalin polypeptides or nucleic acids, ii) detecting the level, or any other detectable property, of placentalin in the sample, and iii) comparing the resulting signal with those obtained from control samples indicative of a benign, a low-grade, an intermediate-grade and a high-grade malignant tumor. The level of the signal negatively correlates with the malignancy of the tumor. 25 Preferably, this method is directed to diagnose the malignancy of tumors such as, e.g., gliomas, amelanotic melanomas, prostate tumors and lung tumors.

[0516] Still another embodiment of the present invention relates to a method of diagnosing the risk for thrombotic deposition comprising the steps of: i) contacting a biological sample with a compound that specifically binds to placentalin polypeptides or nucleic acids, ii) detecting the level, or 30 any other detectable property, of placentalin in the sample, and iii) comparing the resulting signal with those obtained from control samples indicative of a patient presenting no risk for thrombotic

deposition. Preferably, such methods are directed to diagnose hypercoagulable syndromes, e.g., atherosclerosis, ischemia, stroke, myocardial infarction, coronary artery disease.

[0517] Preferably, the biological sample is either a body fluid or tissue sample that has been purified or partially purified to ease detection. Particularly preferred compounds that specifically bind 5 to placentalin polypeptides are the antibodies described above. Preferably, the antibodies are labeled and detection of the signal may be carried out using immunohistological and immunofluorescing processes that are well-known to those skilled in the art. Alternatively, the antibodies that specifically bind to placentalin polypeptides are not labeled and detection of the signal occurs indirectly by using labeled secondary antibodies. Particularly preferred compounds that specifically bind to placentalin 10 polynucleotides are polynucleotide sequences that are complementary to SEQ ID NO:25 or part thereof. Detection of the signal may be carried out using PCR, RT-PCR techniques, or quantitative PCR techniques.

[0518] Still another embodiment relates to a kit for determining the malignancy of a tumor or the risk for thrombotic deposition. Such kits comprise at least one binding partner specific for 15 placentalin polypeptides or nucleic acids coupled to a detectable marker, control samples, and reagents for performing the reaction for detecting the marker. Preferably, the kit is directed to diagnosing the malignancy of a given tumor, e.g., gliomas, amelanotic melanomas, prostate tumors and lung tumors. Alternatively, the kit is directed to monitor the progress of a therapy administered to treat malignant tumors.

20 [0519] Another preferred embodiment relates to substances that increase placentalin expression and to a method of screening for substances that increase placentalin expression comprising the steps of: i) contacting a cell with a test substance; and ii) comparing placentalin expression in the cell after exposure to the test substance to that of an unexposed control cell. Preferably, the test substance modifies the expression of placentalin in a specific cell type while not in others. Most 25 preferably, the test substance modifies placentalin expression specifically in endothelial cells or in neoplastic cells.

[0520] A further embodiment of the present invention is directed to substances that increase placentalin activity and to a method of screening for such substances. This method comprises the steps of: i) contacting a placentalin polypeptide with a test substance, ii) determining placentalin activity, 30 and iii) comparing placentalin activity after exposure to that of an unexposed control sample. Preferably, placentalin activity is studied in an endothelial cell or in a neoplastic cell. Placentalin activity can be monitored by studying plasmin inhibition and/or radiolabeled matrix degradation as

described in Rao *et al.* (Biochem Biophys Res Commun. 276:1286-94 (2000)), which disclosure is incorporated herein in its entirety. Substances that increase placentalin expression or activity are defined as placentalin agonists.

[0521] An embodiment of the present invention relates to methods of using the polypeptides and the polynucleotides of the present invention to reduce blood coagulation in a mammal. Any compositions and methods containing a placentalin polypeptide or a placentalin agonist can be used. Preferably, such methods for reducing blood coagulation are directed to treat a patient presenting a risk for thrombotic deposition. Most preferably, such methods are directed to treat a patient suffering from thrombotic diseases and/or hypercoagulable syndromes, e.g., arteriosclerosis, acute ischemia, stroke, myocardial infarction, coronary artery disease, and thrombotic complications associated with transplantation. A preferred embodiment is a method to reduce blood coagulation comprising the step of introducing a placentalin polypeptide or a protease-inhibiting fragment thereof in a physiologically acceptable composition to the bloodstream of an individual. Delivery to the bloodstream can be direct (e.g., injection to a vein or artery) or indirect (e.g., inhalant or oral delivery), as discussed in detail herein.

[0522] In a further embodiment, the present invention provides a method of promoting wound healing by preventing tissue buildup using compositions comprising a placentalin polypeptide or a placentalin agonist. Such a method comprises the step of: contacting said wound with an effective amount of the composition comprising a placentalin polypeptide or a placentalin agonist, wherein the wound includes, but is not limited to, myocardial infarction, skin wound, arterial wall injury, stents and sutures.

[0523] Another embodiment of the present invention relates to methods of using the polypeptides and the agonists of the present invention to prevent or reduce ECM degradation in a mammal. Preferably, such methods for preventing or reducing ECM degradation are directed to protect the ECM from degradation by malignant cells, thus blocking the invasion and spread of malignant tumors. Most preferably, a composition comprising a placentalin polypeptide is administered to an individual suffering from abnormal or undesirable cell proliferation, such as, e.g., tumor growth, endothelial cell proliferation, and angiogenesis related to tumor growth. Tumors that can be treated with the compositions of the present invention include, but are not limited to, gliomas, amelanotic melanomas, prostate tumors and lung tumors.

[0524] Still another embodiment of the invention relates to the inhibition and/or attenuation of proteases produced by pathogenic microorganisms. This embodiment relates to a method comprising

the step of administering a composition comprising placentalin or a biologically active fragment thereof, or a compound that increases the expression or activity of placentalin, to an individual for preventing and/or treating parasitic infections. Placentalin polypeptides or compounds that increase the expression or activity of placentalin may be administered alone or in combination with other agents. Alternatively, this method may be directed to prevent parasitic infections of mammalian cell cultures. It has been shown that protease inhibitors can prevent dissemination of virus, protozoa, bacteria or fungi in the host organism. Accordingly, the polypeptides of the present invention may be used to prevent or to treat, e.g., coccidiosis, staphylococcal infection, infection by the influenza virus, *P. gingivalis* or *T. denticola*, and invasive pulmonary aspergillosis.

10 [0525] In one embodiment, the pharmaceutical compositions for use in prevention of ECM degradation or inhibition of blood coagulation comprise an effective amount of a placentalin polypeptide, placentalin polynucleotide, or substance that increases placentalin expression or activity in mixture with a pharmaceutically acceptable diluent or carrier. An effective amount of the compositions of the present invention may vary according to factors such as the disease, the disease state, age, sex, and weight of the mammal. Dose regimen may be adjusted to provide the optimum therapeutic response. In a preferred embodiment, a composition comprising a substance that increases placentalin expression or activity and a physiologically acceptable carrier is directly introduced to the bloodstream of an individual by injection.

20 [0526] In a preferred embodiment, a placentalin polypeptide is fused to a targeting molecule specific for the tissue of interest. Preferably, the targeted placentalin polypeptide is comprised in a pharmaceutical composition for use in prevention of ECM degradation associated with tumor growth. Most preferably, the placentalin polypeptide is linked to a ligand or an antibody that recognizes a receptor or an antigen specifically expressed on tumor cells by any technique well-known to those skilled in the art. A large number of tumor-associated antigens have been described in the scientific literature. For example, antigens and techniques described by Wikstrand et al (Cancer Metastasis Rev 18:451-64 (1999)), which disclosure is hereby incorporated by reference in its entirety, may be used. Alternatively, the placentalin polypeptide is linked to a ligand that recognizes the vasculature of a tumor. A preferred method for targeting the placentalin polypeptide to the vasculature of solid tumors is described in U.S. Patent 6,004,554, which disclosure is incorporated by reference in its entirety.

30 [0527] Another embodiment is directed to a method of using placentalin polynucleotides for reducing blood coagulation comprising the steps of: i) constructing a recombinant molecule comprising a nucleic acid sequence encoding a placentalin polypeptide that allows expression of placentalin or part thereof under given physiological conditions, and ii) introducing this recombinant

molecules into a cell, a mammal or a human. Recombinant molecules comprising a nucleic acid sequence encoding a placentalin polypeptide may be directly introduced into cells or tissues *in vivo* using delivery vehicles such as retroviral vectors, adenoviral vectors and DNA virus vectors. They may also be introduced into cells *in vivo* using physiological techniques such as microinjection and 5 electroporation or chemical methods such as coprecipitation and incorporation of DNA into liposomes. Recombinant molecules may also be delivered in the form of an aerosol or by lavage. The placentalin polynucleotides may also be applied extracellularly such as by direct injection into cells. Preferably, placentalin polynucleotides are introduced into endothelial cells or in blood vessel cells. Most preferably, when placentalin polynucleotides are directed to treating a patient presenting a risk for 10 thrombotic deposition, the method described in PCT application WO 99/02171, which disclosure is incorporated by reference in its entirety, can be used.

Protein of SEQ ID NO:28 (Internal designation Clone 176355_117-005-2-0-H11-F)

[0528] The cDNA of Clone 176355_117-005-2-0-H11-F (SEQ ID NO:27) encodes NAAR protein of SEQ ID NO:28, comprising the amino acid sequence:

15 [0529] MASMAAVLTWALALLSAFSATQARKGFWDYFSQTSGDKGRVEQIHQQKMA
REPATLKDSLEQDLNNMNKFLEKLRPLSGSEAPRLPQDPVGMRRQLQEELEEVKARLQPYMA
EAHELGVWNLEGLRQQLKPYTMDLMEQVALRVQELQEQLRVVGEDTKAQLLGGVDEAWA
LLQGLQSRVVHHTGRFKELFHPYAESLVSGIGRHVQELHRSVAPHAPASPARLSRCVQVLSRK
LTLKAKALHARIQQNLDQLREELSRAFAGTGTEEGAGPDPQMLSEEVRQRLQAFRQDTYLQI
20 AAFTRAIDQETEEVQQQLAPPPPQHSAAPEFQQTDSGKVLSKLQARLDDLWEDITHSLHDQG
HSHLGDP. Accordingly, it will be appreciated that all characteristics and uses of the polypeptides of SEQ ID NO:28 described throughout the present application also pertain to the polypeptides encoded by the nucleic acids included in Clone 176355_117-005-2-0-H11-F. In addition, it will be appreciated that all characteristics and uses of the polynucleotides of SEQ ID NO:27 described throughout the
25 present application also pertain to the nucleic acids included in Clone 176355_117-005-2-0-H11-F. A preferred embodiment of the invention is directed toward the compositions of SEQ ID NO:27, SEQ ID NO:28 and Clone 176355_117-005-2-0-H11-F. Also preferred are polypeptide fragments having a biological activity as described herein and the polynucleotides encoding the fragments.

[0530] The protein of the invention displays an Apolipoprotein motif:

30 [0531] MAAVLTWALALLSAFSATQARKGFWDYFSQTSGDKGRVEQIHQQKMAREPA
TLKDSLEQDLNNMNKFLEKLRPLSGSEAPRLPQDPVGMRRQLQEELEEVKARLQPYMAEAHE
LGVWNLEGLRQQLKPYTMDLMEQVALRVQELQEQLRVVGEDTKAQLLGGVDEAWALLQG

LQSRVHHTGRFKELFHPYAESLVSGIGRHVQELHRSVAPHAPASPARLSRCVQLSRKLTAK
AKALHARIQQNLD. Accordingly, an embodiment of the present invention comprises the amino acids of the apolipoprotein motif and the polynucleotides encoding the same.

[0532] Lipoprotein particles such as HDL and LDL contain characteristic apolipoproteins that 5 are responsible for targeting them to certain tissues and for activating enzymes required for the trafficking of the lipid fraction of the lipoprotein, including cholesterol. NAAR is a member of the apolipoprotein family.

[0533] Apolipoproteins are the protein component of lipoprotein particles and are responsible for binding to receptors on cell membranes and directing the lipoprotein particles to their intended site 10 of metabolism. These high molecular weight particles are primarily responsible for lipid transport (triglycerides and cholesterol in the form of cholesteryl esters) through the plasma. Lipoprotein particles include chylomicrons and chylomicrons remnant particles, very low density lipoprotein (VLDL), intermediate density lipoprotein (IDL), low density lipoprotein (LDL), and high density lipoprotein (HDL). Elevated levels of lipoprotein particles have been positively correlated with 15 atherosclerosis and a reduction in plasma lipoprotein concentration is correlated with a reduced risk of atherosclerosis. In addition, individual apolipoproteins have unique functions such as the formation of specific associations with lipoprotein particles of distinct classes. Some apolipoproteins act as ligands controlling the interaction of lipoproteins with cell surface receptors, while others functions as cofactors for essential enzymes in lipid metabolism.

20 [0534] NAAR is associated with the chylomicron and HDL fraction of blood. NAAR synthesis by the small intestine increases markedly after the ingestion of lipid with the result being a marked increase in NAAR output in mesenteric lymph. During secretion from small intestine epithelial cells, the pre-NAAR signal peptide is cleaved. The mature NAAR polypeptide is secreted into the lymph as a major constituent of newly synthesized triglyceride-rich lipoprotein as well as the 25 HDL fraction of blood. NAAR is involved in the biogenesis and/or metabolism of intestinal triglyceride-rich lipoproteins. This increase in biosynthesis and secretion of NAAR by the small intestine is triggered by the formation and secretion of intestinal chylomicrons. NAAR levels are regulated by diet and are up-regulated by high fat meals and down-regulated by leptins. NAAR in mesenteric lymph after a lipid meal suppresses food intake, suggesting that NAAR also acts as a 30 satiety factor that circulates in the blood after fat feeding.

[0535] NAAR plays a role in triglyceride-rich lipoprotein metabolism, in reverse cholesterol transport, and in facilitating CETP (Cholesterol Ester Transfer Protein) activity. As a result, NAAR is

responsible for part of the inter-individual variability in blood cholesterol response to changes in dietary fat/cholesterol intake. Moreover, NAAR may be effectively used in place of HDL particles to prevent the development of atherosclerosis. Also, NAAR plays a significant role in the pathophysiology of diabetes. Levels of NAAR are correlated with glycemic control in young type I diabetes (IDDM) patients and non-insulin-dependent diabetes melitus (NIDDM) patients.

[0536] NAAR plays a major role in lipid metabolism and its related disorders. Diseases involving lipid metabolism include, but are not limited to, obesity, obesity-related disorders such as atherosclerosis, cardiovascular disorders such as coronary heart disease, neurodegenerative disorders such as Alzheimer's disease or dementia, coronary artery disease, mitochondriopathies, 10 hyperlipidemia, familial combined hyperlipidemia (FCHL), hypercholesterolemia, obesity-related arteriosclerosis, obesity-related insulin resistance, obesity related hypertension, microangiopathic lesions resulting from obesity-related Type II diabetes, ocular lesions caused by microangiopathy in obese individuals with type II diabetes, and renal lesions caused by microangiopathy in obese individuals with type II diabetes.

15 [0537] An embodiment of the invention is directed to a composition comprising a NAAR polypeptide sequence of SEQ ID NO:28.

[0538] A further embodiment of the invention is directed to a composition comprising a NAAR polypeptide fragment having biological activity.

20 [0539] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence of SEQ ID NO:27 encoding a NAAR polypeptide.

[0540] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence encoding a NAAR polypeptide fragment having biological activity

[0541] NAAR polypeptides bind free fatty acids (FFAs), triglycerides, and cholesterol, which are associated with atherosclerosis and lipid metabolism-related diseases. A preferred embodiment of 25 the invention is a method comprising the step of: binding a NAAR polypeptide or fragment thereof with an FFA, triglyceride, or cholesterol molecule under conditions that allow binding of NAAR polypeptide to said molecule. In a preferred embodiment, NAAR polypeptide is used to purify FFAs, triglycerides, and cholesterol. In such method, NAAR polypeptide is preferably covalently or non-covalently attached to a solid matrix and allowed to bind FFAs, triglycerides, and cholesterol using 30 techniques well known in the art. This method comprises the steps of: i) washing the solid matrix to get rid of contaminants, ii) eluting the particle of interest using more stringent conditions.

[0542] Additional aspects of this embodiment include methods of using NAAR polypeptide to detect and quantify FFAs, triglycerides, and cholesterol using techniques common in the art. This method comprises the steps of: i) obtaining a biological sample suspected of containing FFAs, triglycerides, or cholesterol; ii) contacting said sample with a NAAR polypeptide or fragment thereof under conditions suitable for binding of NAAR; and detecting the presence or absence of FFAs, triglycerides, and cholesterol by detecting NAAR. Preferably, NAAR polypeptide or fragment thereof is covalently attached to a detectable compound. Alternatively, a detectable NAAR-specific antibody or fragment thereof may be used to detect NAAR. This embodiment is useful, for example, as a diagnostic tool for detecting plasma levels of FFAs, triglycerides, and cholesterol.

10 [0543] In another embodiment, the invention is directed to a method of detecting NAAR polypeptide in a biological sample, said method comprising the steps of: i) contacting a biological sample with an antibody or antibody fragment that specifically binds NAAR polypeptide; and ii) detecting the antigen-antibody complex formed. The antibody or antibody fragment may be monoclonal or polyclonal. In addition, the antibody or antibody fragment may be primarily or 15 secondarily labeled by a detectable compound (e.g., radioactive, fluorescent, luminescent, or enzymatic) common in the art.

[0544] In some embodiments, the invention also concerns a diagnostic kit for detecting *in vitro* the presence of NAAR polypeptide. This kit comprises: a polyclonal or monoclonal antibody or fragment thereof that specifically binds a NAAR polypeptide; and optionally, ii) a reagent allowing the 20 detection of the antigen-antibody complexes formed. Preferably, the antibody or antibody fragment is detectably labeled. Such labels include fluorescent, luminescent, and radioactive compounds, as well as enzymatic substrates. The optional reagent may provide a detectable signal and either bind to the antibody or react with the label on such antibody. NAAR antibodies may be used to diagnose liver-related disorders (e.g., hepatitis, cirrhosis, hepatoma, and FHP), lipid metabolism related disorders 25 such as diabetes, and atherosclerosis. To detect such disorders, an appropriate biological sample (serum, for example) can be tested to determine the level of NAAR being produced (U.S. Patent No. 6,027,935, which disclosure is incorporated by reference in its entirety). (U.S. Patent No. 6,027,935, which disclosure is incorporated by reference in its entirety).

[0545] An embodiment of the invention provides for a method of screening test substances for 30 modulators of NAAR expression. This method comprises the steps of: i) contacting a cell with a test substance; and ii) comparing NAAR expression in the cell after exposure to the test substance to that of an unexposed control cell. NAAR expression is determined by methods common to the art or included herein, by detecting NAAR polynucleotides or polypeptides. An example of this method

comprises the steps of: i) culturing two equivalent cell samples; ii) adding a test substance to one of the cultures and not the other; iii) harvesting both cultures at a specified time; iv) purifying the mRNA from each sample of cells; v) comparing the level of NAAR mRNA in each sample by Northern blot, RTPCR, or another method common to the art. The invention provides for design and use of specific 5 polynucleotide probes and primers, as discussed herein. An additional example comprises the steps of: i) having two equivalent cultures of cells; ii) adding a test substance to one of the cultures and not the other; iii) harvesting both cultures; iv) purifying the protein from each sample of cells; v) comparing the level of NAAR polypeptides in each sample by Western blot, immunohistochemistry, or another method common to the art. The invention provides for design and use of specific antibodies and 10 antibody fragments, as discussed herein.

[0546] In another embodiment, a NAAR polypeptide or a fragment thereof, may be used to screen for compounds that activate or inhibit NAAR activity. This method comprises the steps of: i) contacting a NAAR polypeptide or fragment thereof with a test substance and ii) monitoring NAAR activity. NAAR binds to bind FFAs, triglycerides, and cholesterol. Thus, NAAR activity may be 15 monitored upon addition of the test substance by competitive binding assays with FFAs, triglycerides, or cholesterol. In this aspect of the invention, a NAAR polypeptide or fragment thereof may be free in solution, affixed to a solid support, recombinantly expressed on or chemically attached to a cell surface, or located intracellularly. The formation of binding complexes between NAAR polypeptide and the compound being tested, may be measured by methods well known to those skilled in the art, 20 such as the BIACore (Upsala, Sweden). Another technique provides for high throughput screening of compounds having suitable binding affinity to the protein of the invention as described in published PCT application WO84/03564, and incorporated herein by reference in its entirety. Test substances that decrease NAAR expression or activity are defined as inhibitors or antagonists of NAAR. Test substances that increase NAAR expression or activity are defined as activators or agonists. Agents 25 which modulate the expression or activity of the NAAR of the subject invention include, but are not limited to antisense oligonucleotides, ribozymes, and antibodies. These agents may be made and used according to methods well known in the art.

[0547] In another embodiment of the invention, the NAAR polypeptide is used to bind FFAs, triglycerides, and cholesterol *in vivo* and remove these molecules from the bloodstream. This method 30 comprises the step of: introducing an effective amount of NAAR polypeptide or fragment thereof to the bloodstream of an individual. In this embodiment, the NAAR polypeptide may further be expressed as a fusion protein with a polypeptide signal specifying excretion from the body. A preferred method of delivering NAAR polypeptides or biologically active fragments thereof to an

individual includes direct, intravenous injection of said polypeptides or fragments in a physiologically acceptable solution (e.g., pH-buffered isotonic saline solutions, pH-buffered isotonic saline solutions modified by addition of viscous elements such as glycerol). Alternatively, NAAR polynucleotides may be introduced to express NAAR polypeptides in the bloodstream. This method comprises the 5 steps of: i) constructing a recombinant viral vector corresponding to a portion of the genome of an adenovirus capable of infecting a cell operatively linked to the nucleotide sequence of the invention and a regulatory sequence directing its expression; ii) delivery of an effective amount of the recombinant adenoviral vector to an individual with or at risk of atherosclerosis, diabetes, or other lipid metabolism disorders.

10 [0548] NAAR polypeptides, fragments thereof, or NAAR agonists may be used to prevent or treat atherosclerosis or arterial lipoprotein deposits of FFAs, triglycerides, and cholesterol as determined by common medical techniques, such as determining the familial predisposition of the individual for these disorders or actual analysis of the serum lipoprotein levels in the individual. Alternatively, NAAR polypeptides, fragments thereof, or NAAR agonists may be delivered to an 15 individual at risk of or suffering from diabetes or another lipid metabolism disorder.

[0549] In another embodiment, a NAAR polypeptide, fragment thereof, or NAAR agonist is used in a physiologically acceptable composition to treat septic shock or conditions associated with elevated serum levels of lipopolysaccharide (U.S. Pat Nos. 5,932,536 and 5,948,756, which disclosures are hereby incorporated by reference in their entireties).

20 [0550] In another embodiment, a NAAR polypeptide, fragment thereof, or NAAR agonist is applied to treatment and prevention of obesity, as NAAR peptides have appetite suppressive activities (U.S. Patent 5840688, which disclosure is hereby incorporated by reference in its entirety). This method comprises the step of introducing a NAAR polypeptide, fragment thereof, or NAAR agonist in a physiologically acceptable solution to an individual. Preferably, the individual is overweight or has 25 a body mass index (BMI) of 25 or greater.

[0551] Another embodiment of the subject invention provides compositions and methods of selectively modulating the activity of the protein of the invention. Modulation of NAAR activity would allow for the successful treatment and management of liver-related and lipid metabolism associated disorders.

30 [0552] Another embodiment of the invention relates to composition and methods using polynucleotide sequences encoding the protein of the invention or part thereof to establish transgenic model animals (*D. melanogaster*, *M. musculus*), by any method familiar to those skilled in the art. By

modulating *in vivo* the expression of the transgene with drugs or modifier genes (activator or suppressor genes), animal models can be developed that mimic human hormone-dependent disorders such as cancers. These animal models would thus allow the identification of potential therapeutic agents for treatment of the disorders. In addition, recombinant cell lines derived from these transgenic 5 animals may be used for similar approaches *ex vivo*.

Protein of SEQ ID NO:29 (internal designation Clone 222588_116-094-1-0-H2-F)

[0553] The cDNA of Clone 222588_116-094-1-0-H2-F (SEQ ID NO:29) encodes Neurexinal protein of SEQ ID NO:30, comprising the amino acid sequence:

[0554] MTSGSKCPSTDGKEEYIATFKGSEYFCYDLSQNPIQSSSDEITLSFKTLQRNG
10 LMLHTGKSADYVNLALKNGAVSLVINLGSafeALVEPVNGKFNDNAWHDVKVTRNLRQV
TISVDGILTTGYTQEDYTMGSDDFFYVGGSPTADLPGSPIQHESTFAEDPMFQSQTAQL.
Accordingly, it will be appreciated that all characteristics and uses of the polypeptides of SEQ ID NO:30 described throughout the present application also pertain to the polypeptides encoded by the nucleic acids included in Clone 222588_116-094-1-0-H2-F. In addition, it will be appreciated that all 15 characteristics and uses of the polynucleotides of SEQ ID NO:29 described throughout the present application also pertain to the nucleic acids included in Clone 222588_116-094-1-0-H2-F. A preferred embodiment of the invention is directed toward the compositions of SEQ ID NO:29, SEQ ID NO:30 and Clone 222588_116-094-1-0-H2-F. Also preferred are polypeptide fragments having a biological activity as described herein and the polynucleotides encoding the fragments.

20 [0555] The protein of SEQ ID NO:29, Neurexinal, is a splice variant of Neurexin I-alpha protein (Genbank accession number AB035356). The protein of the invention displays a G-Laminin motif: FKTlQRNGlMLHTGKSADYVNLALKNGAVSLVINLGSafeALVEPVNGKFNDN
AWHDVKVTRNLRQVTISVDGILTTGYTQEDYTMGSDDFFYVGGSPTADLP. Accordingly,
some embodiments of the present invention comprise the amino acids of the G-Laminin motif and the 25 polynucleotides encoding the same.

[0556] Neurexins are neuronal cell-surface proteins. In addition to a function as cell-adhesion molecules, neurexins serve as signaling receptors. Ligands for alpha-neurexins, called neuroxophilins, are similar to peptide hormones. Neurexins can also serve as receptors for latrotoxins. Neurexin splice variants contribute to target innervation and the transition from neurite outgrowth to 30 synaptogenesis or during synaptic plasticity and regeneration. Neurexins are associated with mediating neuronal processes and signaling exocytosis. Thus, these proteins play a role in neuronal development, neuronal abnormalities (e.g., seizures), and cancer.

[0557] Neurexinal, one neurexin isoform, is expressed in sympathetic neurons during target innervation and relative expression levels of splice variants change during differentiation of individual neurons. Neurexinal is a membrane protein detectable only in neurons. Neurexinal expression peaks postnatally. Neurexinal plays a role as a cell adhesion molecule and a signal transduction receptor.

5 [0558] Several ligands to Neurexinal, with putative functions in signaling and cell adhesion have been identified. Neurexinal binds to alpha-latrotoxin with high affinity only in the presence of Ca²⁺. Alpha-latrotoxin is a large peptide neurotoxin from black widow spider venom that causes massive neurotransmitter release when applied to nerve terminals. Alpha-latrotoxin inserts into the membrane and forms a transmembrane ion pore after anchoring to either of its nerve terminal
10 receptors. In addition, Neurexophilin binds tightly to Neurexinal and alpha-Neurexin but not beta-Neurexins. Neurexophilin is synthesized at high levels only in inhibitory interneurons and functions as a signaling molecule. Dystroglycan is another physiological ligand for Neurexinal and mediates cell adhesion between brain cells. Dystroglycan, a cell surface protein, links the intracellular cytoskeleton to the extracellular matrix. Impairment of this linkage is instrumental in the pathogenesis
15 of muscular dystrophies.

[0559] An embodiment of the invention is directed to a composition comprising a novel Neurexinal polypeptide sequence of SEQ ID NO:30.

[0560] A further embodiment of the invention is directed to a composition comprising a Neurexinal polypeptide fragment having biological activity of binding alpha-latrotoxin, neurexophilin,
20 or dystroglycan.

[0561] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence of SEQ ID NO:29 encoding a Neurexinal polypeptide.

[0562] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence encoding a Neurexinal polypeptide fragment having biological activity of
25 binding alpha-latrotoxin, neurexophilin, or dystroglycan..

[0563] A preferred aspect of the invention is a host cell recombinant for polynucleotides encoding a Neurexinal polypeptide or a biologically active fragment thereof. An additional preferred aspect is a host cell recombinant for polynucleotides capable of directing Neurexinal expression. Preferably, the polynucleotides capable of directing Neurexinal expression are located in the 5'
30 regulatory region of the Neurexinal gene. Further preferably, these polynucleotides are located within 500 base pairs of the Neurexinal coding region. These polynucleotides preferably comprise a promoter sequence. Techniques known in the art for introducing polynucleotide sequences to

endogenous sequences are described in U.S. Patent 5641670 and PCT WO9629411, which disclosures are hereby incorporated by reference in their entireties. Neurexinal protein produced by said host cell may be used for in vitro detection and purification methods as well as diagnosis and in vivo applications.

- 5 [0564] Neurexinal polypeptides bind alpha-latrotoxin, neurexophilin and dystroglycan. A preferred embodiment of the invention is a method comprising the step of: binding a Neurexinal polypeptide or fragment thereof with an alpha-latrotoxin, neurexophilin or dystroglycan molecule under conditions that allow binding of Neurexinal polypeptide to said molecule. In a preferred embodiment, Neurexinal polypeptide is used to purify alpha-latrotoxin, neurexophilin or dystroglycan.
- 10 10 In such method, Neurexinal polypeptide is preferably covalently or non-covalently attached to a solid matrix and allowed to bind alpha-latrotoxin, neurexophilin or dystroglycan using techniques well known in the art. This method comprises the steps of: i) washing the solid matrix to get rid of contaminants, ii) eluting the particle of interest using more stringent conditions. Additional aspects of this embodiment include methods of using Neurexinal polypeptide to detect and quantify alpha-
- 15 latrotoxin, neurexophilin and dystroglycan using techniques common in the art. This method comprises the steps of: i) obtaining a biological sample suspected of alpha-latrotoxin, neurexophilin and dystroglycan; ii) contacting said sample with a Neurexinal polypeptide or fragment thereof under conditions suitable for binding of Neurexinal; iii) and detecting the presence or absence of alpha-latrotoxin, neurexophilin and dystroglycan by detecting Neurexinal. Preferably, Neurexinal
- 20 polypeptide or fragment thereof is covalently attached to a detectable compound. Alternatively, a detectable Neurexinal -specific antibody or fragment thereof may be used to detect Neurexinal. This embodiment is useful, for example, as a diagnostic tool for detecting neuronal levels of alpha-latrotoxin, neurexophilin and dystroglycan. A further aspect of the present invention involves the isolation, purification and characterization of Neurexinal. Specifically, the present invention extends
- 25 25 to a novel neuronal receptor which is a regulator of neurotransmitter release, and thus mediates alpha-latrotoxin toxicity in the presence of calcium.

[0565] An embodiment of the present invention relates to compositions comprising Neurexinal polypeptides. The method of producing Neurexinal polypeptides comprises the steps of: i) transfecting a mammalian host cell with a recombinant expression vector comprising a polynucleotide of the present invention, and ii) purifying the produced protein. Neurexinal may be isolated according to any technique known in the art or disclosed herein. Preferably, an antibody directed against Neurexinal or a fragment thereof, is bound to a chromatographic support to form an affinity chromatography column.

[0566] In another embodiment, the protein of the invention may be used to target heterologous compounds (polypeptides or polynucleotides) to the neurons. For instance, a chimeric protein composed of Neurexinal or a fragment thereof recombinantly or chemically fused to a compound, protein, or polynucleotide of interest would allow specific delivery to cells that express 5 ligands for neurexinal, such as dystroglycan. Dystroglycan is expressed in neurons of the cerebral cortex, hippocampus, olfactory bulb, basal ganglia, thalamus, and hypothalamus. A method of targeting these cell populations would be valuable for treatment of a number of neuronal and mood disorders such as epilepsy, stress disorder, schizophrenia, Huntingtons, Alzheimers and Parkinsons disease. In addition, the hypothalamus of obese individuals often expresses lower than normal levels 10 of leptin receptor. Thus, introduction of leptin receptor constructs or agonists to the hypothalamus with heterologous Neurexinal may be an effective treatment for obesity.

[0567] A preferred embodiment of the invention relates to compositions and methods using the protein of the invention or fragment thereof to label neurons in order to visualize any change in number, topology or morphology. Such methods can be used to map neuronal tissue damage 15 associated with stroke; neurodegenerative conditions such as schizophrenia, Alzheimer's disease, epilepsy, stress disorder, Huntington's disease, and Parkinson's disease; and peripheral neuromuscular diseases such as myasthenia gravis and AIDS-related complex. For example, the protein of the invention or a fragment thereof may be used to generate specific antibodies, which would in turn allow the visualization of neurons, by methods well known to those skilled in the art. Preferably, the 20 antibody is labeled with any detectable moiety including, but not limited to, a fluorescent label, a radioactive atom, a paramagnetic ion, biotin, a chemiluminescent label or a label which can be detected through a secondary enzymatic or binding step. In a similar fashion, antibodies raised against Neurexinal may be used to identify particular cell types as Neurexinal is specifically expressed only in neurons. Alternatively, quantitative analysis or detection of the protein of the invention, or of nucleic 25 acids encoding Neurexinal, can be carried out by any other technique known to those skilled in the art. The invention further provides a method of diagnosing neurological disease or conditions, including but not limited to those listed above.

[0568] In another embodiment, the invention relates to a method of using compounds that bind a Neurexinal polypeptide or fragment thereof. Such method is directed toward diagnosis of 30 disorders related to abnormal Neurexinal levels including neurodegenerative conditions such as schizophrenia, Alzheimer's disease, epilepsy, stress disorder, Huntington's disease, and Parkinson's disease. This method comprises the step of contacting a compound that specifically binds Neurexinal with a biological sample. Preferred compounds include Neurexinal specific antibodies and antigen

binding fragments thereof. Further preferred compounds include those labeled with a detectable label, such as a fluorescent, radioactive, or enzymatic tag. This method may rely on various diagnostic techniques, including immunoassays.

[0569] In another embodiment, the present invention includes an assay system which may be 5 prepared in the form of a test kit for the quantitative analysis of the extent of the presence of Neurexinal, or to identify compounds that may mimic or block its activity. The system or test kit comprises: a component that specifically binds Neurexinal and optionally, a means for detecting said component. Preferred components are Neurexinal-specific antibodies and antigen-binding fragments thereof. Preferably, the antibody or fragment is detectably labeled with, for example, a fluorescent, 10 luminescent, or radioactive compound. Alternatively, the Neurexinal binding component may be labeled with a detectably enzymatic substrate and the optional means is the enzyme required for detection.

[0570] It is a further object of the present invention to provide a method and associated assay system for screening substances that are effective in either mimicking the activity or combating the 15 adverse effects of Neurexinal in mammals.

[0571] A preferred embodiment of the invention is a method of screening for test compounds that increase or decrease Neurexinal expression. This method comprises the steps of: i) contacting a cell with a test substance, ii) detecting and quantifying Neurexinal expression levels, and iii) comparing Neurexinal expression in an exposed cell to that of an unexposed control cell. Preferably, 20 Neurexinal expression is studied in neurons. Methods of detecting Neurexinal expression may be directed either to Neurexinal transcripts (e.g., Northern blotting, in situ hybridization, and other nucleotide detecting methods) or Neurexinal polypeptides (e.g., Western blotting or immunohistological methods). The expression level of Neurexinal can be decreased in a biological sample, for example using antisense molecules.

[0572] A preferred embodiment of the invention is a method of screening for compounds that 25 increase or decrease Neurexinal activity. This method comprises the steps of: i) contacting a cell with a test substance, ii) detecting and quantifying Neurexinal activity levels, and iii) comparing Neurexinal activity in an exposed cell to that of an unexposed control cell. Preferably, Neurexinal activity is detected by determining binding of Neurexinal to alpha-latrotoxin, neurexophilin or dystroglycan. 30 Binding may be tested by competitive binding assays with detectably-labeled proteins from the above list.

[0573] Neurexinal activators or agonists are defined as compounds that increase Neurexinal activity or expression. Neurexinal inhibitors or antagonists are defined as compounds that decrease Neurexinal activity or expression.

[0574] In yet a further embodiment, the invention is drawn to antagonists of the activity of Neurexinal. For example, the antagonist may be used to counteract alpha-latrotoxin and used as an antivenom for black widow spider bites. A Neurexinal antagonist may further be used or to treat disorders associated with abnormally high norepinephrine levels, such as Pheochromocytoma, attention-deficit hyperactivity disorder (ADHD), idiopathic torsion dystonia, and monoamine oxidase A deficiency. A method of using a Neurexinal inhibitor comprises the step of: administering a Neurexinal inhibitor in a physiologically acceptable composition to an individual. Physiologically acceptable compositions and carriers are discussed herein. Appropriate treatment will vary depending on the individual situation and may be determined by one skilled in the art.

[0575] Neurexinal activators may be used to modulate and/or reverse the degeneration of nerve terminals, to modulate synaptic transmission, or to treat other pathologies related to abnormally low levels of Neurexinal expression. A method of using a Neurexinal activator comprises the step of: administering a Neurexinal activator in a physiologically acceptably solution to an individual. Physiologically acceptable compositions and carriers are discussed herein. Appropriate treatment will vary depending on the individual situation and may be determined by one skilled in the art.

Protein of SEQ ID NO:32 (internal designation Clone 119033_105-066-4-0-F10-F)

[0576] The cDNA of Clone 119033_105-066-4-0-F10-F (SEQ ID NO:31) encodes NPIASY protein of SEQ ID NO:32, comprising the amino acid sequence: MAAELVEAKNMVMSFRVSDLQMLLGFGVGRSKGLKHELVTRALQLVQFDCTPELFKKIKEL YETRYAKKNSEPAPQPHRPLDPLTMHSTYDRAGAVPRTPLAGPNIDYPVLYGKYLNGLGRLP AKTLKPEVRLVKLPFFNMLDELLKPTELVPQNNEKLQESPCIFALTPRQVELIRNSRELQPGVK AVQVVLRICYSDTSCPQEDQYPPNIAVKVNHSYCSVPGYYPSNKPGVEPKRPCRPINLTHLMY LSSATNRITVTWGNYGKSYSVALYLVRQLTSSELLQRLKTIGVKHPELCKALVKEKLRLDPDS EIATTGVRVSLICPLVKMRLSVP CRAETCAHLQCFDAVFYLQMNEKKPTWMCPVCDKPAPYD QLIIDGLLSKILSECEDADEIEYLVDGSWCPIRAEKELSCSPQGAILVLGPSDANGLLPAPS VNG SGALGSTGGGPVGSMENGKPGADVVDLTLDSSSSSEEEEEEEDEDEEGPRPKRRCPFQ KGLVPAC. Accordingly, it will be appreciated that all characteristics and uses of the polypeptides of SEQ ID NO:32 described throughout the present application also pertain to the polypeptides encoded by the nucleic acids comprising the human cDNA in Clone 119033_105-066-4-0-F10-F. In addition,

it will be appreciated that all characteristics and uses of the polynucleotides of SEQ ID NO:31 described throughout the present application also pertain to the nucleic acids included in Clone 119033_105-066-4-0-F10-F. A preferred embodiment of the invention is directed toward the compositions of SEQ ID NO:31, SEQ ID NO:32 and Clone 119033_105-066-4-0-F10-F. Also 5 preferred are polypeptide fragments having a biological activity as described herein and the polynucleotides encoding the fragments.

[0577] The protein of SEQ ID NO:32, NPIASY, is a novel polymorphic variant of Protein Inhibitor of Activated STAT Protein (PIASY) (accession numbers O75926). The protein of the invention displays a SAP domain: VMSFRVSDLQMLLGFVGRSKSGLKHELVTRALQLV. In 10 addition, NPIASY displays a MIZ Zinc finger motif:

[0578] VSLICPLVKMRLSVP CRAETCAHLQCFDAVFY LQMNEKETCAHLQCFDAVFY LQMNEK.

[0579] Activation of early response genes by interferons (IFNs) and other cytokines requires tyrosine phosphorylation of a family of transcription factors termed signal transducer and activator of 15 transcription (STAT) proteins. STAT proteins relay signals from activated cell surface receptors directly to the nucleus and play a critical role in gene induction by a variety of hematopoietic cytokines and hormones, such as the interleukin 6 (IL6) family of cytokines, epidermal growth factor, and leptin. All STAT proteins contain a DNA binding domain, a Src homology 2 domain and a transactivation domain necessary for transcriptional activation of target gene expression. Janus kinases (JAK) are 20 cytoplasmic protein kinases which, upon cytokine signaling, phosphorylate STAT proteins on a specific tyrosine residue. Tyrosine phosphorylated STAT proteins dimerize through specific reciprocal SH2-phosphotyrosine interactions and translocate from the cytoplasm to the nucleus. Once in the nucleus, STAT proteins stimulate the transcription of specific target genes by binding to response elements in their promoters.

25 [0580] The binding of proteins from the Protein Inhibitor of Activated STAT (PAIS) family inhibits the STAT signaling pathway. The PIAS-STAT interaction depends on the tyrosine phosphorylation of STAT proteins.

[0581] NPIASY is localized in the nucleus. Upon Interferon stimulation, NPIASY becomes associated with STAT1. An LXXLL coregulator sequence at the NH₂-terminus of NPIASY, although 30 not involved in NPIASY-STAT1 interaction, is required for repression of STAT1-mediated gene activation. NPIASY is a transcriptional corepressor of STAT1. Also, NPIASY acts as a potent inhibitor of transcriptional activity of androgen receptor in prostate cancer cells. NPIASY binds to

androgen receptor but does not affect the DNA biding activity of androgen receptor. The NH2 terminal LXXLL signature motif of NPIASY, although not required for NPIASY-androgen receptor interaction, is essential for the inhibitory activity of NPIASY. In addition, NPIASY binds to p53 and inhibits p53-mediated transactivation. However, NPIASY does not affect p53-mediated apoptosis.

5 NPIASY regulates p53-mediated functions and directs p53 into a transactivation-independent mode of apoptosis.

[0582] An embodiment of the invention is directed to a composition comprising a NPIASY polypeptide sequence of SEQ ID NO:32.

[0583] A further embodiment of the invention is directed to a composition comprising a 10 NPIASY polypeptide fragment having a biological activity of binding to STAT, p53, or Androgen Receptor.

[0584] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence of SEQ ID NO:31 encoding a NPIASY polypeptide.

[0589] A further embodiment of the invention is directed to a composition comprising a 15 polynucleotide sequence encoding a NPIASY fragment having a biological activity of binding to STAT, p53, or Androgen Receptor.

[0590] NPIASY polypeptides bind to STAT, androgen receptor, and p53, which are often associated with cancers including T cell leukemias, breast, prostate, colorectal, pancreatic, lung, and esophageal cancers. A preferred embodiment of the invention is a method comprising the step of 20 binding a NPIASY polypeptide, or fragment thereof with STAT, androgen receptor or p53 molecules under conditions that allow binding of NPIASY polypeptide to said molecules. This method may be directed to in vitro purification and detection methods as well as in vivo methods to inhibit STAT, p53, or Androgen Receptor. Such methods are discussed herein.

[0591] In a preferred embodiment, NPIASY polypeptide is used to purify STAT, androgen 25 receptor or p53. In such method, NPIASY polypeptide is preferably covalently or non-covalently attached to a solid matrix and allowed to bind STAT, androgen receptor or p53 using techniques well known in the art. This method comprises the steps of: i) washing the solid matrix to get rid of contaminants, ii) eluting the particle of interest using more stringent conditions. Additional aspects of this embodiment include methods of using NPIASY polypeptide to detect and quantify STAT, 30 androgen receptor, or p53 using techniques common in the art. This method comprises the step of: i) obtaining a biological sample suspected of containing STAT, androgen receptor, or p53; ii) contacting said sample with a NPIASY polypeptide or fragment thereof under conditions suitable for binding of

NPIASY; and iii) detecting the presence of STAT, androgen receptor, or p53 by detecting NPIASY. Preferably, NPIASY polypeptide or fragment thereof is covalently attached to a detectable compound. Alternatively, a detectable NPIASY-specific antibody or fragment thereof may be used to detect NPIASY. This embodiment is useful, for example, as a diagnostic tool for detecting risk of p53, 5 STAT, or Androgen Receptor related cancers such as T cell leukemias, breast, prostate, colorectal, pancreatic, lung, and esophageal cancers. In addition, this method may be used to diagnose disorders associated with a low level of the above proteins including muscle atrophy and sexual development defects.

[0592] A preferred aspect of the invention is a host cell recombinant for polynucleotides 10 encoding a NPIASY polypeptide or a biologically active fragment thereof. An additional preferred aspect is a host cell recombinant for polynucleotides capable of directing NPIASY expression. Preferably, the polynucleotides capable of directing NPIASY expression are located in the 5' regulatory region of the NPIASY gene. Further preferably, these polynucleotides are located within 500 base pairs of the LIRION coding region. These polynucleotides preferably comprise a promoter 15 sequence. Techniques known in the art for introducing polynucleotide sequences to endogenous sequences are described in U.S. Patent 5641670 and PCT WO9629411, which disclosures are hereby incorporated by reference in their entireties. NPIASY protein purified from said host cell may be used for in vitro methods of detection and purification, as well as diagnostic and in vivo applications.

[0593] An embodiment of the present invention relates to a method of producing 20 compositions comprising NPIASY polypeptides. The method of producing NPIASY polypeptides comprises the steps of: i) transfecting a host cell with a recombinant expression vector comprising a polynucleotide of the present invention, and ii) purifying the produced protein. NPIASY may be purified by any technique known in the art or disclosed herein. Preferably, an antibody directed against NPIASY or part thereof, preferably, an antibody directed against the C-terminal sequence of 25 NPIASY polypeptide, may be bound to a chromatographic support to form an affinity chromatography column. Alternatively, NPIASY polypeptides may be produced by a method comprising the step of: i) transfecting a host cell with a polynucleotide capable of directing NPIASY expression. Preferably, the polynucleotides capable of directing NPIASY expression are located in the 5' regulatory region of the NPIASY gene. Further preferably, these polynucleotides are located within 500 base pairs of the 30 NPIASY coding region. These polynucleotides preferably comprise a promoter sequence, as discussed above.

[0594] In another embodiment, the invention is directed to a method of detecting NPIASY polypeptide in a biological sample, said method comprising the steps of: i) contacting a biological

sample with an antibody or antibody fragment that specifically binds NPIASY polypeptide; and ii) detecting the antigen-antibody complex formed. The antibody or antibody fragment may be monoclonal or polyclonal. In addition, the antibody or antibody fragment may be primarily or secondarily labeled by a detectable compound common in the art (e.g., radioactive, fluorescent, 5 luminescent, or enzymatic). This method may be applied to diagnosis of NPIASY-related disorders. For example, an abnormally low level of NPIASY may indicate a high risk of p53, STAT, or Androgen Receptor related cancers such as T cell leukemias, breast, prostate, colorectal, pancreatic, lung, and esophageal cancers. An abnormally high level of NPIASY may indicate a risk of growth and 10 cell proliferation defects associated with the above proteins such as incomplete sexual development and muscle atrophy.

[0595] In some embodiments, the invention also concerns a diagnostic kit for detecting the presence of NPIASY polypeptide. This kit comprises: i) a polyclonal or monoclonal antibody or fragment thereof that specifically binds a NPIASY polypeptide; and optionally, ii) a reagent allowing the detection of the antigen-antibody complexes formed. Preferably, the antibody, or antibody 15 fragment is detectably labeled. Such labels include fluorescent, luminescent, and radioactive compounds, as well as enzymatic substrates. The optional reagent may provide a detectable signal that either binds to the antibody or reacts with the label on the antibody. NPIASY antibodies may be used to diagnose any of the disorders listed above. To diagnose such disorders, an appropriate biological sample can be tested to determine the level of NPIASY being produced.

[0596] In another embodiment, the invention provides methods for regulating an IFN-associated immune response mediated by STAT. The immune response includes an anti-viral or anti-tumor response mediated by IFN secreted by B and/or T cells. This method comprises the step of introducing a polynucleotide construct comprising polynucleotides encoding a NPIASY polypeptide to an IFN-responsive cell. Preferably, the polynucleotide construct comprises an expression control 20 element operably linked to the NPIASY-encoding polynucleotide. Suitable compositions of polypeptides or polynucleotides of the present invention are useful as a method of treatment of IFN-related pathologies such as inflammation.

[0597] An embodiment of the invention provides for a method of screening test substances for modulators of NPIASY expression. This method comprises the steps of: i) contacting a cell with a test 30 substance; ii) detecting and quantifying NPIASY expression; and ii) comparing NPIASY expression in the cell after exposure to the test substance to that of an unexposed control cell. NPIASY expression is determined by methods common to the art or included herein, by detecting NPIASY polynucleotides or polypeptides. An example of this method comprises the steps of: i) culturing two equivalent cell

samples; ii) adding a test substance to one of the cultures and not the other; iii) harvesting both cultures at a specified time; iv) purifying the mRNA from each sample of cells; v) comparing the level of NPIASY mRNA in each sample by Northern blot, RTPCR, or another method common to the art. The invention provides for design and use of specific polynucleotide probes and primers, as discussed 5 herein. An additional example comprises the steps of: i) having two equivalent cultures of cells; ii) adding a test substance to one of the cultures and not the other; iii) harvesting both cultures; iv) purifying the protein from each sample of cells; v) comparing the level of NPIASY polypeptides in each sample by Western blot, immunohistochemistry, or another method common to the art. The invention provides for design and use of specific antibodies and antibody fragments, as discussed 10 herein.

[0598] In another embodiment, a NPIASY polypeptide or a fragment thereof, may be used to screen for compounds that activate or inhibit NPIASY activity. This method comprises the steps of: i) contacting a NPIASY polypeptide or fragment thereof with a test substance and ii) monitoring NPIASY activity. NPIASY binds to STAT, androgen receptor, and p53. Thus, NPIASY activity may 15 be monitored upon addition of the test substance by competitive binding assays with either STAT, androgen receptor, or p53. In this aspect of the invention, a NPIASY polypeptide or fragment thereof may be free in solution, affixed to a solid support, recombinantly expressed on or chemically attached to a cell surface, or located intracellularly. The formation of binding complexes between NPIASY polypeptide and the compound being tested, may be measured by methods well known to those skilled 20 in the art, such as the BIACore (Upsala, Sweden). Another technique provides for high throughput screening of compounds having suitable binding affinity to the protein of the invention as described in published PCT application WO84/03564, and incorporated herein by reference in its entirety. Test substances that decrease NPIASY expression or activity are defined as inhibitors or antagonists of NPIASY. Test substances that increase NPIASY expression or activity are defined as activators or 25 agonists. Antagonists of the polypeptides or polynucleotides of the present invention are used to stimulate, promote, or facilitate progression through the cell cycle, such as in the cellular regeneration of terminally differentiated cardiac myocytes or tissues, e.g., striated muscle myocytes. For example, this could allow restoration of damaged myocardium after cardiac injury, myocardial infarction, myocarditis, cardiomyopathy, trauma, as a consequence of cardial surgery, etc., or repletion of striated 30 muscle exhausted by muscular dystrophy. Agonists of polypeptides or polynucleotides of the present invention are useful in treatment of pathologies such as but not limited to hyperproliferative diseases such as cancer (e.g., leukemia, lymphoma, breast cancer, colon cancer, prostate cancer) coronary artery disease, pulmonary vascular, obstructive disease, either primary or as a feature of

Eisenmenger's syndrome, and other disorders of abnormal cellular proliferation. Agents which modulate the expression or activity of NPIASY include, but are not limited to antisense oligonucleotides, ribozymes, and antibodies. These agents may be made and used according to methods well known in the art.

5 [0599] In a further embodiment, expression of the polypeptides encoded by the nucleic acids is expected to prevent, ameliorate, or lessen the cell cycle defect of the host cell, or to restore normal cell cycle progression of the host cell. Whether provided via nucleic acid or polypeptides delivered directly to cells, the therapeutic formulations of the invention can also be used as adjuncts to other forms of therapy, including but not limited to chemotherapy, and radiation therapy.

10 [0600] In another preferred embodiment, the protein of the invention can be used to modulate and/or characterize fertility, including for the treatment or diagnosis of infertility, and for contraception. NPIASY inactivates Androgen receptor which is required for the activation of genes essential for spermatogenesis. Thus, over-expression or activation of NPIASY can be used to repress Androgen receptor transcription activity and thereby inhibit fertility. For example, for contraception, 15 the expression of genes for spermatogenesis can be artificially disrupted, for example by increasing the protein level using polynucleotides encoding the protein, using the protein itself, or using activators of protein expression or activity. Alternatively, for infertility, the protein level can be decreased using inhibitors such as antisense oligonucleotides, antibodies, dominant negative forms of the protein, and using heterologous compounds that inhibit protein expression or activity. Similarly, the cause of 20 infertility in many patients can be detected by detecting the level of expression of the present protein, where an abnormal level of activity or expression of the protein indicates that a cause of infertility involves the NPIASY repression. Such a diagnosis would also point to methods of treating the infertility, e.g. by increasing or decreasing the expression or activation of the present protein.

[0601] Another embodiment of the invention relates to composition and methods using 25 polynucleotide sequences encoding the protein of the invention or a fragment thereof to establish transgenic model animals (*D. melanogaster*, *M. musculus*), by any method familiar to those skilled in the art. By modulating *in vivo* the expression of the transgene with drugs or modifier genes (activator or suppressor genes), animal models can be developed that mimic human hormone-dependent disorders such as cancers. These animal models would thus allow the identification of potential 30 therapeutic agents for treatment of the disorders. In addition, recombinant cell lines derived from these transgenic animals may be used for similar approaches *ex vivo*.

Protein of SEQ ID NO:34 (Internal designation 125402_105-074-4-0-F3-F)

[0602] The cDNA of 125402_105-074-4-0-F3-F (SEQ ID NO:33) encodes protein vCRTL-1 of SEQ ID NO:34, comprising the amino acid sequence:

MLLSLTLSLVLLGSSWGCGIPAICKPALSFSQRIVNGENAVLGSPWPQVSLQDSSDFHFC

GGSLISQSWVVTAAHCNVSPGRHFVVLGEYDRSSNAEPLQVLSVSRAITHPSWNSTMNN

5 DVTLLKLASPAQYTTRISPVCLASSNEALTEGLCTTGWGRLSGVGNVTPARLQQVALP
LTVVNCRQYWGSITDSMICAGGAGASSCQGDGGPLVCQKGNTWVLIGIVSWGKNCNV
RAPAVYTRVSKFSTWINQVIAYN. Accordingly, it will be appreciated that all characteristics and uses of polypeptides of SEQ ID NO:34 described throughout the present application also pertain to the polypeptides encoded by the nucleic acids included in 125402_105-074-4-0-F3-F. In addition, it will
10 be appreciated that all characteristics and uses of the polynucleotides of SEQ ID NO:33 described throughout the present application also pertain to the nucleic acids included in 125402_105-074-4-0-F3-F. A preferred embodiment of the invention is directed toward the compositions of SEQ ID NO:33, SEQ ID NO:34, and 125402_105-074-4-0-F3-F. Also preferred are polypeptide fragments having a biological activity as described herein and the polynucleotides encoding the fragments.

15 [0603] The cDNA of SEQ ID NO:33 is a novel polymorphic variant of the human chymotrypsin-like protease CTRL-1 encoded by a gene located on chromosome 16, specifically at position 16q22.1. The 264 amino-acid protein of vCRTL-1 contains a signal peptide (MLLLSLTLSLVLLGSSWG) and a serine protease domain and belongs to the serine protease family.

[0604] Proteases are key components in a broad range of biological pathways and can be
20 classified into four groups according to their catalytic mechanisms: the serine, cysteine (thiol), aspartic (carboxyl), and metalloproteases. Chymotrypsins are members of a family of enzymes known as serine proteases, so named because they utilize an activated serine residue in their substrate-binding site to catalyze the hydrolysis of certain peptide bonds. The pancreatic subfamily of proteases comprises trypsin, chymotrypsin, kallikrein, and elastases. Chymotrypsins are the most abundant
25 among the pancreatic proteases and may represent 10-20% of the total protein synthesized by the exocrine pancreas. vCRTL-1 is a digestive enzyme expressed and synthesized by pancreatic acinar cells as an inactive zymogen precursor that must be cleaved to form an active protease. vCRTL-1 displays chymotrypsin- and elastase-2-like activities and hydrolyzes the amide bonds of substrates having tyrosine, phenylalanine, or leucine residues at the P1 position, and preferably a proline in the
30 P2 position. The secretion of vCRTL-1 can be induced by the presence of protease inhibitors in the pancreas or intestinal fluid (in particular duodenal fluid).

[0605] An embodiment of the invention is directed to a composition comprising a vCRTL-1 polypeptide sequence of SEQ ID NO:34.

[0606] A further embodiment of the invention is directed to a composition comprising a vCRTL-1 polypeptide fragment having a biological activity of protein substrate binding or serine protease activity.
5

[0607] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence of SEQ ID NO:33 encoding a vCRTL-1 polypeptide.

[0608] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence encoding a vCRTL-1 polypeptide fragment having a biological activity of
10 protein substrate binding or serine protease activity.

[0609] A further embodiment of the invention is directed to a composition comprising an antibody directed against the vCRTL-1 polypeptide sequence of SEQ ID NO: 34 or a vCRTL-1 polypeptide fragment having a biological activity of protein substrate binding or serine protease activity. Preferably, the antibody specifically binds to vCRTL-1 but not to CTRL-1. Even more
15 preferably, the antibody recognizes the SLQDSSDFHF amino-acid sequence.

[0610] An additional preferred embodiment of the invention is a method of binding vCRTL-1 polypeptides with a vCRTL-1-specific antibody or vCRTL-1-binding fragment thereof. This method comprises the step of contacting a vCRTL-1 polypeptide with a vCRTL-1-binding antibody or vCRTL-1-binding fragment thereof under conditions that allow binding. This method may be applied
20 to detection and purification of vCRTL-1. These aspects are discussed in detail herein.

[0611] Another embodiment relates to a method of producing a recombinant vCRTL-1 polypeptide comprising the steps of: i) transfecting a host cell with an appropriate expression vector comprising a polynucleotide encoding vCRTL-1 and ii) purifying the produced protein. The host cells are cultured under conditions whereby the nucleic acid sequence coding for this particular protein is
25 expressed. After a suitable amount of time for the product to accumulate, the protein is purified from the host cells or medium surrounding the cells. Introduction of an expression vector incorporating a nucleic acid sequence coding for vCRTL-1 into a host cell can be performed in a variety of ways, such as but not limited to calcium or lithium chloride treatment, electroporation, and lipofection. Also in this embodiment, a milk animal can be used to produce vCRTL-1 in the milk, thereby generating a
30 significant amount of this particular protein after purification. Any type of animal that produces enough milk can be used in this aim such as, but not limited to, sheep, goat, and cow. These animals can be generated with any method of targeting overexpression of vCRTL-1 in the milk. The protein of

the invention may be purified using any technique well-known to those skilled in the art including those disclosed in the US Patent 6,268,487, which disclosure is hereby incorporated by reference in its entirety. Preferably, an antibody or antigen-binding fragment directed against vCRTL-1 or a fragment thereof may be bound to a chromatographic support to form an affinity chromatography column. Even 5 more preferably, the antibody binds specifically to vCRTL-1 but not to CTRL-1.

[0612] A preferred embodiment of the invention is a method of using vCRTL-1 polypeptide or a biologically active fragment thereof to proteolytically cleave a substrate protein. This method comprises the step of contacting vCRTL-1 with a protein substrate under conditions that allow vCRTL-1 protease activity. This method is directed to both *in vitro* and *in vivo* uses of the protein that 10 include biochemical applications and prevention and treatment of vCRTL-1-related disorders such as chronic and acute pancreatitis and cystic fibrosis.

[0613] In another embodiment, the present invention is to provide a method of effectively inducing vCRTL-1 proteolytic activity to digest a wide range of proteins in a biological sample. This method can be performed either *in vitro* or *in vivo* and comprises the step of contacting an effective 15 amount of vCRTL-1 polypeptide or a biologically active fragment thereof with a biological sample under conditions that allow vCRTL-1 proteolytic activity. For example, such a method can be used *in vitro* for removing proteins in a DNA preparation or for removing enzymes after any enzymatic reaction. Such a composition can be used alone or as a "cocktail" with other proteases. Such protease cocktails are widely used in assays to degrade proteins susceptible of interfering with an assay to be 20 performed.

[0614] In another embodiment, vCRTL-1 or fragment thereof can be used in combination with a detergent for the removal of stains having a protein component, similar to the use of proteases described in U.S. Patent 5,599,400, which disclosure is incorporated by reference in its entirety. The composition can contain known detergent constituents such as surfactants, foam enhancers and fillers. 25 vCRTL-1 polypeptides can be included in a detergent composition or can be combined with other constituents at the time of use as an additive. The detergent additive can be formulated as a liquid, powder, granulate or slurry.

[0615] A further embodiment of the present invention is directed to substances that increase or decrease vCRTL-1 expression, and to a method of screening for such substances comprising the 30 steps of: i) contacting a cell with a test substance, ii) quantifying and detecting vCRTL-1 expression levels, and iii) comparing vCRTL-1 expression in an exposed cell to that of an unexposed control cell. vCRTL-1 expression may be determined by any specific nucleotide or protein detection techniques

common to the art (e.g., Northern or Western blotting) Preferably, vCRTL-1 expression is studied in pancreatic cells. The expression level of vCRTL-1 can be decreased in a biological sample, for example using antisense molecules.

[0616] A further embodiment of the present invention is directed to substances that increase 5 or decrease vCRTL-1 activity, and to a method of screening for such substances comprising the steps of: i) contacting a cell with a test substance and ii) detecting vCRTL-1 activity, and iii) comparing vCRTL-1 proteolytic activity in an exposed cell to that of an unexposed control cell. Preferably, vCRTL-1 expression is studied in pancreatic cells. The activity of vCRTL-1 may be inhibited using direct or indirect inhibitor molecules or antagonistic antibodies.

10 [0617] Activators or agonists of vCRTL-1 are substances that increase the activity or expression of vCRTL-1. Inhibitors or antagonists of vCRTL-1 are substances that decrease the activity or expression of vCRTL-1.

[0618] The present invention also provides animal models generated by modulating the expression or activity of the present protein in one or more tissues of the animal and preferably in 15 pancreas. Such animals are useful for a number of purposes, because they represent an *in vivo* assay method for testing candidate molecules potentially useful for the treatment of various pathophysiological aspects of diseases specifically related to the activity of vCRTL-1 such as acute or chronic pancreatitis. Study of the phenotype of such models can also allow the identification of additional human equivalent diseases caused by or linked with vCRTL-1 deficiency. These animals 20 can be generated with any method of targeting overexpression or inactivation of vCRTL-1. Such models are extremely useful, e.g. in the assessment of candidate therapies and drugs for the treatment of such diseases and conditions.

[0619] In another embodiment, the current invention is used to diagnose diseases or disorders associated with abnormal vCRTL activity. In particular, it is useful in diagnosing patients with 25 exocrine pancreatic insufficiency which results in a dramatic reduction of vCRTL-1 activity. Examples of such diseases and disorders include, but are not limited to, chronic or acute pancreatitis and cystic fibrosis. In addition, vCRTL is associated with metastatic tumor activity, in particular, that of pancreatic tumors. Thus, higher than normal levels of vCRTL may be indicative of metastatic 30 activity. The method includes the steps of: i) contacting a fluid or tissue sample with a compound capable of selectively binding vCRTL-1 polypeptide or polynucleotide and ii) detecting the level of vCRTL-1 polypeptide or polynucleotide in the sample. Preferably, a difference in the level or other property in the sample relative to in a control sample indicates the presence of the disorder or a

propensity for developing disorder. Preferably, samples to be analyzed include, but are not limited to, intestinal fluids and in particular duodenal juice, homogenates of pancreatic tissue, urine and plasma. More preferably, the fluid or tissue sample is obtained from an individual suspected of suffering from the disorder, or at risk of developing the disorder. For example, a polyclonal or monoclonal antibody 5 or any antigen binding active fragment thereof or a nucleic acid probe may be used. Preferably, the antibodies used in this embodiment are specifically directed against the variant vCRTL-1 polypeptides and do not recognize CRTL-1 polypeptides. Detection can be carried out directly or indirectly with known immunohistological or immunofluorescent processes. vCRTL-1 binding molecules may be directly labeled with generally known compounds, including, but not limited to, enzymes such as 10 alkaline phosphatase and peroxidase and fluorescent dyes such as FITC, rhodamine, and Texas-Red. However, labeling can also occur indirectly by using secondary reagents. Also in this embodiment, diagnosis may be facilitated by the identification of the variant vCRTL-1 using well known PCR or RT-PCR techniques and in particular in with “real-time” PCR system. Alternatively, using such a method, the present invention provides a tool to correlate modulations in the expression of vCRTL-1 15 with certain pathologies. Thus, the present invention provides a novel candidate gene for such conditions.

[0620] In a further embodiment, the invention includes a test kit useful for the quantification of the amount of vCRTL-1 in a biological sample. The level of vCRTL-1 polypeptide is useful in evaluating the degree of severity of pancreatic disorders that have been clinically ascertained and also 20 represents a useful marker in examining the pathophysiology and possible treatment modalities in the animal model of acute pancreatitis. In addition, vCRTL-1 level may be useful in the estimation of tumor invasiveness and metastasis ability, preferably in the case of pancreatic cancers. In this embodiment, the application of such assays can be used to monitor the progress of therapy administered to treat these or other conditions. Further, the assays can be used as a measure of 25 toxicity, for example, during clinical testing of new drugs to assess the impact on tissue degradation. Thus, the assays may be applied in any situation wherein vCRTL-1 may be used as an index of the condition, treatment, or effect of substances directly administered to the subject. Thus, the condition of a patient can be monitored continuously and the quantified amount of vCRTL-1 in the pathological sample can be compared with the amount quantified in a biological sample of a normal individual or 30 with the previous analysis of the same patient. In this embodiment, this marker can be measured effectively in intestinal fluids and in particular in duodenal juice, homogenates of pancreatic tissue, urine and plasma, by any suitable method, including immunoassays. The kit comprises at least one immunological binding partner (e.g., a monoclonal or polyclonal antibody specific for vCRTL-1) and

optionally, a secondary detectable marker. Preferably, the antibody or antigen-binding fragment thereof used in the kit is specifically directed against the variant vCRTL-1 polypeptides and not to CTRL-1 polypeptides.

[0621] A further embodiment of the present invention is to provide novel methods and compositions useful for treating individuals at risk of developing or suffering from a disorder associated with insufficient secretion of vCRTL-1. These methods and compositions are useful for the treatment of exocrine pancreatic deficiency characterized by steatorrhoea such as chronic or acute pancreatitis and cystic fibrosis. This method comprises the step of: administering an effective amount of vCRTL-1 polypeptide, a biologically active fragment thereof, or agonist to an individual with or at risk of one of the listed conditions. Preferably, vCRTL-1 or agonist compositions are delivered in combination with a physiologically acceptable carrier, such as a saline solution or other physiologically buffer. The particular amount of the compositions of the invention that will be administered to the individual for any particular condition will depend on the clinical condition of the patient, and other factors such as the weight, age, and route of delivery. Such compositions may be administered by any suitable route including, but not limited to, oral, aerosol, intravenous, intramuscular, intraperitoneal, or subcutaneous routes. These compositions can comprise the protein of the invention, and, optionally, one or more other types of protease, or any other compound of interest. This co-administration may be by simultaneous administration or by separate or sequential administrations. All of these additional components may be either obtained from natural sources or produced by recombinant genetic engineering techniques and/or chemical modification.

[0622] A further embodiment of the present invention is to provide novel methods and compositions useful for treating individuals at risk of developing or suffering from a disorder associated with higher than normal levels of vCRTL-1. This method is useful for the treatment of metastatic cell growth, in particular, pancreatic cancers. This method comprises the step of administering an effective amount of a vCRTL-1 antagonist to an individual with or at risk of developing a metastatic cell growth. Preferably, vCRTL-1 antagonist compositions are delivered in combination with a physiologically acceptable carrier, such as a saline solution or other physiologically buffer. The particular amount of the compositions of the invention that will be administered to the individual for any particular condition will depend on the clinical condition of the patient, and other factors such as the weight, age, and route of delivery. Such compositions may be administered by any suitable route including, but not limited to, oral, aerosol, intravenous, intramuscular, intraperitoneal, or subcutaneous routes.

[0623] The cDNA of Clone 107640 105-036-2-0-H3-F (SEQ ID NO:35) encodes the 447 amino acid LIRION protein of SEQ ID NO:36 comprising the amino acid sequence:

[0624] MIPTFTALLCLGLSLGPRTHMQAGPLPKPTLWAEPGSVISWGNSVTIWCQGTL
EAREYRLDKEESPAPWDRQNPLEPKNKARFSIPSMTEDYAGRYRCYYRSPVGWSQPSDPLEL
5 VMTGAYSKPTLSALPSPLVTSEKSVTLLCQSRSPMDTFLLIKERAHPLLHLRSEHGAQQHQAE
EFPMSPVTSVHGHTYRCFSSHGFSHYLLSHPSDPLELIVSGSLEDPRPSPTRSVSTAAGPEDQPL
MPTGSVPHSGLRRHWEVLIGVLVVSILLSSLLFLLQHWRQGKHRTLAQRQADFQRPPGAA
EPEPKDGLLQRSSPAADVQGENFCAAVKDTQPEDGVEMDTRSPHDDEDPQAVTYAKVKHSR
PRREMASPPSPLSGEFLDTKDRQAEDRQMDTEAAASEAPQDVTYAQLHSFTLRQKATEPPPS
10 QEGASPAEPSVYATLAIH. Accordingly, it will be appreciated that all characteristics and uses of polypeptides of SEQ ID NO:36 described throughout the present application also pertain to the polypeptides encoded by the nucleic acids comprising the human cDNA in Clone 107640 105-036-2-0-H3-F. In addition, it will be appreciated that all characteristics and uses of the polynucleotides of SEQ ID NO:35 described throughout the present application also pertain to the nucleic acids included
15 in 107640 105-036-2-0-H3-F. Also preferred are fragments having a biological activity as described therein and the polynucleotides encoding the fragments.

[0625] LIRION, the protein of SEQ ID NO:36, represents a novel variant form of the ILT-3 receptor precursor (Genbank entry U82979) starting with a 259 amino acid long extracellular domain including a peptide signal sequence (MIPTFTALLCLGLSLGPRTHMQA) and two C2 type immunoglobulin-like-domains, followed by a transmembrane region (VLIGVLVVSILLSSLLFLL) and a 167 amino acid long intracellular domain which contains three Immunoreceptor Tyrosine-based Inhibitory Motifs (ITIMs).

[0626] Immune system cellular activity is controlled by a complex network of cell surface interactions and associated signaling processes. Upon receptor binding, cellular activity is regulated by
25 a balance between activating and inhibitory signals. Many receptors that mediate positive signaling have cytoplasmic tails containing sites of tyrosine phosphorylation known as Immunoreceptor Tyrosine-based Activation Motifs (ITAMs). The inhibitory pathways involve receptors having ITIMs, which provide binding sites for protein tyrosine phosphatases.

[0627] The cytolytic activity of Natural Killer (NK) cells is regulated by a balance between
30 positive signals that initiate cell function and inhibitory signals which prevent the cell activity. Normal cells are protected from NK cells by cell surface expression of MHC class I proteins that are recognized by inhibitory receptors on the NK cells. These receptors, known as Killer Inhibitory

Receptors (KIRs) send a negative signal to the cell upon engagement and down regulate NK cell cytotoxic activity. MHC class I expression is often down-regulated in virally infected and in tumor cells, which renders these cells susceptible to NK cell attack. The KIRs are immunoglobulin superfamily receptors with two ITIMs. These phosphorylated ITIMs recruit tyrosine phosphatases which dephosphorylate molecules in the signal transduction pathway and prevent cell activation.

[0628] LIRION, the protein of the invention, is an inhibitory Immunoglobulin-Like Transcript (ILT) receptor. This class of receptor is closely related to the KIRs and functions in the same manner. LIRION is selectively expressed on myeloid antigen presenting cells (APCs), i.e. dendritic cells, monocytes and macrophages and negatively regulates the functional response of APCs triggered by stimulatory receptors. In addition to its inhibitory function, LIRION is involved in antigen uptake and processing.

[0629] Preferred LIRION polypeptides for use in the methods described below include the polypeptides comprising the amino sequence of:

10 GPLPKPTLWAEPGSVISWGNSVTIWCQGTLEAREYRLDKEESPAPWDRQNPLEPKNKARFSIP
15 SMTEDYAGRHYRCYYRSPVGWSQPSDPLEVMTGAYSKPTLSALPSPLVTSEKSVTLLCQSRSP
 MDTFLLIKERAHPLLHLRSEHGAQQHQAEFPMSPVTSVHGHTYRCFSSHGFSHYLLSHPSDP
 LELIVSGSLEDPRPSPTRSVSTAAGPEDQPLMPTGSVPHSGLRRHWEVLIGVLVVSILLSLLF
 LLLQHWRQGKHRTLAQRQADFQRPPGAAEPEPKDGGLQRRSSPAADVQGENFCAAVKDTQP
 EDGVEMDTRSPHDDEDPQAVTYAKVKHSRPREMASPPSPLSGEFLDTKDRQAEDRQMDTE
20 AAASEAPQDVTVYAQLHSFTLRQKATEPPPSQEGASPAEPSVYATLAIH;

[0630] A polypeptide comprising the amino acid sequence of:

10 GPLPKPTLWAEPGSVISWGNSVTIWCQGTLEAREYRLDKEESPAPWDRQNPLEPKNKARFSIP
15 SMTEDYAGRHYRCYYRSPVGWSQPSDPLEVMTGAYSKPTLSALPSPLVTSEKSVTLLCQSRSP
 MDTFLLIKERAHPLLHLRSEHGAQQHQAEFPMSPVTSVHGHTYRCFSSHGFSHYLLSHPSDP
 LELIVSGSLEDPRPSPTRSVSTAAGPEDQPLMPTGSVPHSGLRRHWE
25 QHWRQGKHRTLAQRQADFQRPPGAAEPEPKDGGLQRRSSPAADVQGENFCAAVKDTQPED
 GVEMDTRSPHDDEDPQAVTYAKVKHSRPREMASPPSPLSGEFLDTKDRQAEDRQMDTEAA
 ASEAPQDVTVYAQLHSFTLRQKATEPPPSQEGASPAEPSVYATLAIH.

[0631] A polypeptide comprising the amino acid sequence of:

10 QHWRQGKHRTLAQRQADFQRPPGAAEPEPKDGGLQRRSSPAADVQGENFCAAVKDTQPED
15 GVEMDTRSPHDDEDPQAVTYAKVKHSRPREMASPPSPLSGEFLDTKDRQAEDRQMDTEAA
 ASEAPQDVTVYAQLHSFTLRQKATEPPPSQEGASPAEPSVYATLAIH.

30 [0632] An embodiment of the invention is directed to a composition comprising a LIRION polypeptide sequence of SEQ ID NO:36.

[0633] A further embodiment of the invention is directed to a composition comprising a LIRION polypeptide fragment having biological activity of interacting with SHP-1, MHC I, MHC I-like molecules (such as CD1, MR1, MIC), MHCII, and complement components (such as C1, C2, C3, C4, C5, and factor B).

5 [0634] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence of SEQ ID NO:35 encoding a LIRION polypeptide.

[0635] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence encoding a LIRION polypeptide fragment having biological activity of interacting with SHP-1, MHC I, MHC I-like molecules (such as CD1, MR1, MIC), MHCII, and 10 complement components (such as C1, C2, C3, C4, C5, and factor B).

[0636] A further embodiment of the invention is directed to a method of screening test substances for modulators of LIRION expression comprising the steps of: i) contacting a cell with a substance to be tested; and ii) comparing LIRION expression in the cell after exposure to the test substance to that of an untreated control cell.

15 [0637] A preferred embodiment relates to an antibody that specifically binds to LIRION polypeptides. Further included are LIRION binding antibody fragments. Preferably, the LIRION-specific antibody or LIRION-binding fragment thereof binds LIRION and not ILT3. Monoclonal, polyclonal, heterologous, and detectably labeled antibodies and antigen-binding fragments are included in the invention. Methods of making such antibodies and fragments thereof are common to 20 the art and included herein.

[0638] An additional preferred embodiment of the invention is a method of binding LIRION polypeptides with a LIRION-specific antibody or LIRION-binding fragment thereof. This method comprises the step of contacting a LIRION polypeptide with a LIRION-binding antibody or LIRION-binding fragment thereof under conditions that allow binding. This method may be applied to 25 detection and purification of LIRION, as well as targeting LIRION-expressing cells. These aspects are discussed in detail herein.

[0639] In one embodiment, a sequence encoding SEQ ID NO:36 bearing G to A, G to A and A to G substitutions at nucleotide positions 447, 705 and 1040 of SEQ ID NO:35 corresponding to positions 137, 223 and 335, and resulting in the substitution of a glycine residue by a glutamic acid at 30 position 137, a glycine by an aspartic acid at position 223 and an asparagine residue by an aspartic acid at position 335, respectively, can be used for DNA genotyping. Genotyping this locus could be of interest, e.g., in DNA fingerprinting for paternity studies or forensic analyses. It could also be used for

genetic association studies, especially in pathologies relating to B cell autoimmune disorders (e.g., rheumatoid arthritis and ulcerative colitis) and antigen presentation disorders (such as bare lymphocyte syndrome)..

[0640] In another embodiment, the polynucleotide sequence of the invention can be used in pharmacogenomic applications in order to aid in the choice of the ideal drug (e.g. an agonist or an antagonist of LIRION), or dosage of a drug, for the treatment of a condition or disease in a patient. For example, in one embodiment, the invention provides a method of genotyping a patient to determine the identity of the nucleotides encoding the amino acids at positions 137, 223 and 335 of LIRION, and administering to the patient a drug or a dosage of the drug that has been established to be preferentially efficacious in those with glutamic acid, aspartic acid and aspartic acid residues at positions 137, 223 and 335 (e.g. because of preferential binding of the drug to the isoform of the protein with glutamic acid, aspartic acid and aspartic acid at these positions). In another embodiment, the patient is genotyped for the nucleotides encoding amino acid positions 137, 223 and 335, a drug is not administered, e.g. because side effects are known to be associated with the administration of the drug to individuals with glutamic acid, aspartic acid and aspartic acid at positions 137, 223 and 335, respectively.

[0641] A preferred aspect of the invention is a host cell recombinant for polynucleotides encoding a LIRION polypeptide or a biologically active fragment thereof. An additional preferred aspect is a host cell recombinant for polynucleotides capable of directing LIRION expression. Preferably, the polynucleotides capable of directing LIRION expression are located in the 5' regulatory region of the LIRION gene. Further preferably, these polynucleotides are located within 500 base pairs of the LIRION coding region. These polynucleotides preferably comprise a promoter sequence. Techniques known in the art for introducing polynucleotide sequences to endogenous sequences are described in U.S. Patent 5641670 and PCT WO9629411, which disclosures are hereby incorporated by reference in their entireties.

[0642] Another embodiment relates to methods of producing LIRION polypeptides. The protein of the invention can be produced in host cells that have been transfected with an appropriate expression vector comprising a nucleic acid sequence coding for LIRION. Introduction of such expression vector to a host can be performed in a variety of ways, including calcium or lithium chloride treatment, electroporation, or lipofection. Any of a wide variety of expression systems can be used to express the recombinant proteins. Suitable expression vehicles include, but are not limited to plasmids, viral particles or baculovirus for insect cells. The expression vehicle can be integrated into the host cell genome. Optionally, an inducible expression vector can be used to achieve controlled

expression of the gene in the host cell. The recombinant protein can be recovered from the host cell and purified by any technique well known to those skilled in the art. Preferably, an antibody directed against the protein of the invention or part thereof can be bound to a chromatographic support to form an affinity chromatography column. Alternatively, LIRION or fragments thereof can be chemically synthesized using solid-phase techniques.

[0643] In another embodiment, the present protein can be used to purify myeloid APCs from a biological sample or from cells grown in vitro. In one such a method a monoclonal or polyclonal antibody directed against the protein of the invention can be used to immunopurify the cells. Methods used to immunopurify cells are well known in the art. For example, the antibody can be fixed on a 10 sepharose column or linked to a solid support. Alternatively, the antibody is fluorescently labeled and cells are purified by FACS.

[0644] In another embodiment, the invention provides compositions and methods for detecting the level of expression of the mRNA encoding the protein of the invention in a mammal, preferably a human. Quantification of mRNA levels of LIRION may be useful for the diagnosis of 15 conditions correlated with abnormal expression of the protein, including autoimmune disorders, or to monitor regulation of LIRION during therapeutic intervention as described herein. Assays for the detection and quantification of mRNA are well known in the art. Preferred method comprises the steps of: isolating RNA from a biological sample, measuring LIRION mRNA level by quantitative RT-PCR, and comparing the expression in the subject sample to that of a control sample. 20 Polynucleotide probes or primers used for the detection of LIRION mRNA by hybridization or PCR amplification may be designed from the polynucleotide of SEQ ID NO:35 by methods well known in the art.

[0645] In another embodiment, the invention relates to methods and compositions for detecting LIRION and quantifying its level of expression in a biological sample. These methods may 25 be useful for the diagnosis of conditions characterized by altered or abnormal LIRION expression including autoimmune (e.g., rheumatoid arthritis, multiple sclerosis, and virally-induced autoimmunity), metabolic (e.g., obesity, hypercholesterolemia, insulin-related disorders), and neurological (e.g., schizophrenia) conditions. A method to quantify the expression of LIRION protein comprises the step of: contacting a biological sample with a LIRION-specific antibody or LIRION 30 specific fragment thereof under conditions that allow antibody-antigen binding. Preferably, the antibody or fragment is detectably labeled. Preferred labels include fluorescent labels (e.g., GFP, FITC, and rhodamine) and secondary labels (e.g., biotin and enzymatic substrates). A preferred method comprises contacting an antibody which specifically binds to LIRION with a biological

sample from a mammalian subject, and determining the level of LIRION in the subject sample compared to a control level of a healthy subject, wherein abnormal level of LIRION in the subject sample indicates that the subject has the disease or is at an elevated risk of developing the disease. The antibody used can be either monoclonal or polyclonal and can be labeled directly or indirectly for 5 quantification of immune complexes by methods well known to those skilled in the art, for example by ELISA or radioimmunoassays. Preferred specific antibodies directed to LIRION are generated using the polypeptide fragments listed above. Diagnostic assays to detect the protein of the invention may rely on a biological sample of a biopsy or in situ assay of cells from organ or tissue sections. Preferred biological samples include whole cells and cellular extracts from organs, tissues, or blood.

10 [0646] An embodiment of the invention provides for a method of screening test substances for modulators of LIRION expression. This method comprises the steps of: i) contacting a cell with a test substance; ii) detecting expression of LIRION; and iii) comparing LIRION expression in the cell after exposure to the test substance to that of an unexposed control cell. LIRION expression is determined by methods common to the art or included herein, by detecting LIRION polynucleotides or 15 polypeptides. An example of this method comprises the steps of: i) culturing two equivalent cell samples; ii) adding a test substance to one of the cultures and not the other; iii) harvesting both cultures at a specified time; iv) purifying the mRNA from each sample of cells; v) detecting the level of LIRION mRNA in each sample. LIRION mRNA may be detected by Northern blot, RT-PCR, or another method common to the art. The invention provides for design and use of specific 20 polynucleotide probes and primers, as discussed herein. An additional example comprises the steps of: i) having two equivalent cultures of cells; ii) adding a test substance to one of the cultures and not the other; iii) harvesting both cultures; iv) purifying the protein from each sample of cells; v) detecting the level of LIRION polypeptides in each sample. LIRION may be detected by enzyme-linked immunoabsorbent assay (ELISA), western blot, radioimmunoassay (RIA), or another method common 25 to the art. The invention provides for design and use of specific antibodies and antibody fragments, as discussed herein.

[0647] A preferred embodiment of the invention provides a method of screening for test substances that bind LIRION polypeptides. This method comprises the steps of: i) contacting a test substance with a LIRION polypeptide or fragment thereof under conditions that allow binding; and ii) 30 detecting the binding of the test substance by methods common to the art (e.g., competitive antibody-based methods such as coimmunoprecipitation and Western blotting). Included in this method are test substances that are conjugated to an antibody, antibody fragment, cell-type specific ligand or a portion

thereof. Substances that bind specifically to LIRION may be used to detect and purify the protein, as well as inhibit or activate its activity.

[0648] A further preferred embodiment of the invention provides a method of screening test substances that bind to LIRION for agonists of LIRION activity, comprising: i) contacting a cell with 5 the substance to be tested; and ii) comparing LIRION biological activity after exposure to the test substance to that of an unexposed control cell. Measure of LIRION biological activity may be assessed indirectly by monitoring signalling events following co-engagement with an activating receptor. An example of an assay to measure LIRION activation comprises the steps of culturing monocytes in vitro, contacting the cells with the test substance, an anti-LIRION antibody and an anti-10 HLA-DR antibody, adding a cross-linking antibody, harvesting the cells, separating cell extracts by SDS-PAGE and detecting protein tyrosine phosphorylation using a specific antiphosphotyrosine antibody. A decrease in tyrosine protein phosphorylation in the test sample compared to the control sample being indicative of an activating effect of the test substance on LIRION. Activation of LIRION can also be assessed by measuring intracellular Ca⁺⁺ mobilization induced upon receptor co-15 ligation as described in Cella et al. (1997, Journal of Experimental Medicine; Vol.185, pp 1743-1751, which disclosure is hereby incorporated by reference in its entirety).

[0649] A further preferred embodiment of the invention provides a method of screening test substances that bind to LIRION for antagonists of LIRION activity. This method comprises the steps of: i) contacting a cell with a test substance; and ii) comparing LIRION biological activity after 20 exposure to the test substance to that of an unexposed control cell. Detection of LIRION biological activity may be detected by measuring protein dephosphorylation following LIRION stimulation as mentioned above. In this case, an increase in protein phosphorylation in the test sample compared to the control sample is indicative of an inhibitory effect of the test substance on LIRION activity.

[0650] Substances that increase LIRION expression or activity (agonists or activators) may be 25 used to decrease APC activity or treat myeloid cancers. Substances that decrease LIRION expression or activity (antagonists or inhibitors) may be used to treat or prevent infectious diseases.

[0651] Another embodiment of the invention is also directed to methods to treat or prevent disorders associated with suppressed immune function such as pathogenic infections. This method comprises the step of introducing a LIRION antagonist to a cell. Preferred cells are those that express 30 LIRION, such as APCs.

[0652] A preferred embodiment of the invention includes methods of increasing LIRION expression or activity by administering a LIRION agonist or polypeptides encoding LIRION to a cell.

These methods may be applied to LIRION-associated conditions, including autoimmune (e.g., osteoarthritis, Crohn's disease, Grave's disease, autoimmune thyroiditis, lupus erythematosus, autoimmune hemolytic anemia, thrombocytopenia, atherosclerosis, osteoporosis, asthma, myasthenia gravis, rheumatoid arthritis, multiple sclerosis, and virally-induced autoimmunity), metabolic (e.g., 5 obesity, hypercholesterolemia, insulin-related disorders), and neurological (e.g., schizophrenia) conditions as well as myeloid cancers. One such method comprises the step of contacting a cell with a LIRION agonist. Preferred agonists include those that affect LIRION expression, such as heterologous promoter sequences. An additional method comprises the step of introducing a polynucleotide encoding a LIRION polypeptide to a cell. One example of this method includes: i) 10 removing a sample of APC cells of an individual; ii) introducing ex vivo a polynucleotide sequence complementary to LIRION polynucleotide sequences to those cells, and iii) reinjecting the recombinant cells into the individual. APCs include monocytes, macrophages and dendritic cells. Expression vectors derived from retroviruses, adenovirus, herpes, or vaccinia viruses may be used for the delivery of the polynucleotide sequences complementary to those of the invention to the targeted 15 cell population. Many methods for introducing vectors into cells or tissues are available and equally suitable for use either in vivo or ex vivo. Delivery by transfection and liposome injections may be achieved using methods which are well known in the art. Pharmaceutical compositions comprising LIRION agonists, antagonists, or oligonucleotides can be administered alone or in combination with at least one other agent, such as stabilizing compound, which may be administered in any sterile 20 biocompatible carrier. They may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intranasal, enteral or rectal means. The compositions may be administered to the subject alone or in combination with other agents, or drugs.

[0653] It is yet a further object of the invention to provide compositions and methods to 25 deliver into a cell a molecule of interest comprising contacting cells expressing LIRION with a molecule joined to a LIRION antibody. Preferably, these methods are to be applied in vivo. A preferred method comprises joining a molecule of interest with a compound that specifically binds the extracellular portion of LIRION; and contacting the chimeric molecule with a cell that express under conditions that trigger LIRION internalization. The targeted cell populations may be naturally 30 expressing the protein of the invention at their cell surface or may be transfected by LIRION encoding sequences prior to the addition of the chimeric antibody. Antibodies to be used in such methods are directed against the extracellular domain of the mature protein, preferably comprising the epitope VTSEKSVT or a LIRION-specific fragment thereof. Molecules to be delivered into the cells can be

fluorescent or radioactive dyes or therapeutically active molecules, including but not limited to chemical drugs, radioactive compounds, peptides or nucleic acids. Methods to link an antibody moiety to a second molecule are well known by those skilled in the art. Alternatively, chimeric proteins with LIRION antibodies may also be obtained by synthesis *in vivo*. Using these methods 5 would be of particular interest in gene therapy, radioimmunosintigraphy (RIS) and radioimmunotherapy (RIT) of specific LIRION-expressing cell types such as APCs.

Protein of SEQ ID NO:38 (Internal designation Clone 588394 160-105-4-0-A11-F)

[0654] The cDNA of Clone 588394 160-105-4-0-A11-F (SEQ ID NO:37) encodes the 383 amino acid SLAMP protein of SEQ ID NO:38 comprising the amino acid sequence:
10 TYWLHSVWVLGFFLSLFLQGLPVRSVDFNRGTDNITVRQGDTAILRCVVEDKNSKVAWLNR
SGIIFAGHDKWSDLPRVELEKRHSLEYSLRIQKVVDVYDEGSYTCVQTQHEPKTSQVYLIVQV
PPKISNISSDVTVNNEGSNVTLCMANGRPEPVITWRHLPTGREFEGEEYLEILGITREQSGKY
ECKAANEVSSADVVKQVKVTVNYPPTITESKSNEATTGRQASLKCEASAVPAPDFEWYRDDTR
INSANGLEIKSTEGQSSLTVTNVTEEHGNYTCVAANKLGVTNASLVLFKRVLPTIPHPIQEIGT
15 TVHFKQKGIFLSESQRGETTKITLNCGNLFLRNLHPTSDQEPQQLWTLCCLLPRKGQHRIYGQC
. Accordingly, it will be appreciated that all characteristics and uses of polypeptides of SEQ ID NO:38 described throughout the present application also pertain to the polypeptides encoded by the nucleic acids comprising the human cDNA in Clone 588394 160-105-4-0-A11-F. In addition, it will be appreciated that all characteristics and uses of the polynucleotides of SEQ ID NO:37 described 20 throughout the present application also pertain to the nucleic acids included in 588394 160-105-4-0-A11-F. Also preferred are fragments having a biological activity as described therein and the polynucleotides encoding the fragments.

[0655] SLAMP, the protein of SEQ ID NO:38, represents a novel splice variant of the limbic system-associated membrane protein (LAMP; Swissprot entry Q13449). SLAMP has a 22 amino acid 25 signal sequence (MRTYWLVHSVWVLGFFLSLFLQ), three C2-type immunoglobulin-like-domains, and a unique 80 amino acid carboxy-terminus. The protein belongs to the IgLON family, a group of neuronal glycoproteins which have been isolated from chicken, rat and human. These include the LAMP protein, Neurotrimin/CEPU-1, OBCAM and kilon/Neurotractin polypeptides. Most of the known members of this family are cell-surface adhesion molecules with a glycosyl 30 phosphatidylinositol anchor at their C-terminus which tethers them to the neuronal plasma membrane. SLAMP, however, lacks the GPI anchor and is thus secreted by neurons in cortical and subcortical regions of the limbic system. SLAMP interacts with itself and other IgLON proteins and is involved in cell-cell recognition, contact-dependent regulation of neurite out-growth and axon guidance of

specific subset of neurons during brain development. SLAMP interacts with LAMP, Neurotrimin and OBCAM, promotes neurite outgrowth of limbic neurons and inhibits neurite outgrowth of non-limbic neurons both *in vivo* and *in vitro*.

[0656] Preferred SLAMP polypeptides for uses in the methods described below include the 5 polypeptides comprising the amino sequence of:

GLPVRSVDFNRGTDNITVRQGDTAILRCVVEDKNSKVAWLNRSGIIFAGHDKWSLDPRVELE
KRHSLEYSLRIQKVDTVYDEGSYTCVQTQHEPKTSQVYLIVQVPPKISNISSDVTVNEGSNVTL
VCMANGRPEPVITWRHLTPTGREFEGEREEYLEILGITREQSGKYECKAANEVSSADVKQVKVT
VNYPPTITESKSNEATTGRQASLKCEASAVPAPDFEWYRDDTRINSANGLEIKSTEGQSSLTVT
10 NVTEEHYGNYTCVAANKLGVTNASLVLFKRVLPTIPHPIQEIGTTVHFKQKGIFLSESQRGETT
KITLNCGNLFLRNLLHPTSDQEPQRLWTLCLLPRKGQHRIYGQC;

[0657] A polypeptide comprising the amino acid sequence of:

GLPVRSVDFNRGTDNITVRQGDTAILRCVVEDKNSKVAWLNRSGIIFAGHDKWSLDPRVELE
KRHSLEYSLRIQKVDTVYDEGSYTCVQTQHEPKTSQVYLIVQVPPKISNISSDVTVNEGSNVTL
15 VCMANGRPEPVITWRHLTPTGREFEGEREEYLEILGITREQSGKYECKAANEVSSADVKQVKVT
VNYPPTITESKSNEATTGRQASLKCEASAVPAPDFEWYRDDTRINSANGLEIKSTEGQSSLTVT
NVTEEHYGNYTCVAANKLGVTNASLVLF;

[0658] A polypeptide comprising the amino acid sequence of:

KRVLPTIPHPIQEIGTTVHFKQKGIFLSESQRGETTKITLNCGNLFLRNLLHPTSDQEPQRLWTLCLLPRKGQHRIYGQC.

[0659] An embodiment of the invention is directed to a composition comprising a SLAMP polypeptide sequence of SEQ ID NO:38.

[0660] A further embodiment of the invention is directed to a composition comprising a SLAMP polypeptide fragment having biological activity of binding to any one of the group consisting 25 of LAMP, Neurotrimin, CEPU-1, CEPU-1-Se, OBCAM, kilon, Neurotractin, and GP55-A.

[0661] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence of SEQ ID NO:37 encoding a SLAMP polypeptide.

[0662] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence encoding a SLAMP polypeptide fragment having biological activity of

binding to any one of the group consisting of LAMP, Neurotrimin, CEPU-1, CEPU-1-Se, OBCAM, kilon, Neurotractin, and GP55-A.

[0663] A further embodiment of the invention is directed to a method of screening test substances for modulators of SLAMP expression comprising the steps of: i) contacting a cell with a substance to be tested; and ii) comparing SLAMP expression in the cell after exposure to the test substance to that of an untreated control cell.

[0664] A preferred aspect of the invention is a host cell recombinant for polynucleotides encoding a SLAMP polypeptide or a biologically active fragment thereof. An additional preferred aspect is a host cell recombinant for polynucleotides capable of directing SLAMP expression. 10 Preferably, the polynucleotides capable of directing SLAMP expression are located in the 5' regulatory region of the SLAMP gene. Further preferably, these polynucleotides are located within 500 base pairs of the SLAMP coding region. These polynucleotides preferably comprise a promoter sequence. Techniques known in the art for introducing polynucleotide sequences to endogenous sequences are described in U.S. Patent 5641670 and PCT WO9629411, which disclosures are hereby 15 incorporated by reference in their entireties.

[0665] One embodiment of the present invention relates to methods of producing SLAMP polypeptides. The protein of the invention can be produced in host cells that have been transfected with an appropriate expression vector comprising a nucleic acid sequence coding for the protein of the invention. Introduction into a host cell of such expression vector for SLAMP can be performed in a 20 variety of ways, including but not limited to calcium or lithium chloride treatment, electroporation, or lipofection. Any of a wide variety of expression systems can be used to provide the recombinant proteins. Suitable expression vehicles include, but are not limited to plasmids, viral particles or baculovirus for insect cells. The expression vehicle can be integrated into the host cell genome. Optionally, an inducible expression vector can be used to achieve tight controlled expression of the 25 gene in the host cell. The recombinant protein can be recovered from the host cell or the cell culture medium and purified by any technique well known to those skilled in the art. Preferably, a SLAMP-specific antibody SLAMP-specific fragment thereof can be bound to a chromatographic support to form an affinity chromatography column. Alternatively, SLAMP or fragments thereof can be chemically synthesized using solid-phase techniques.

30 [0666] In a further embodiment, the invention provides compositions and methods using SLAMP or fragments thereof to label neurons in order to visualize any change in number, topology or morphology of these cells. Preferred neurons are those expressing LAMP, Neurotrimin, CEPU-1,

CEPU-1-Se, OBCAM, kilon, Neurotractin, and GP55-A. These changes may be associated with a central nervous system disorder. SLAMP may be rendered easily detectable by labeling or conjugating with a detectable molecule or antibody.

[0667] In another embodiment, the present protein can be used to purify neuronal cells from a biological sample such as neurons grown in vitro. This method comprises the steps of: i) contacting a cell sample with SLAMP or a biologically active fragment thereof; ii) removing non-binding contaminants; and iii) eluting the remaining cells. Preferred cell samples are samples comprising neuronal cells. Further preferred are cells expressing LAMP, Neurotrimin, CEPU-1, CEPU-1-Se, OBCAM, kilon, Neurotractin, and GP55-A. Preferably, SLAMP is fixed on a sepharose column or covalently or noncovalently linked to a solid support. Alternatively, the neuronal cells can be preincubated with SLAMP and then a monoclonal or polyclonal antibody directed against SLAMP can be used to immunopurify the cells, for example on a column. Alternatively, the SLAMP-specific antibody is fluorescently labeled and cells are purified by FACS.

[0668] In another embodiment, the invention provides compositions and methods for detecting the level of expression of the mRNA encoding the protein of the invention in a mammal, preferably a human. This method comprises the steps of: contacting a SLAMP mRNA or cDNA molecule with a detectable compound under conditions that allow binding. Preferably, the detectable compound is a nucleotide sequence complimentary to SLAMP mRNA. Preferably, SLAMP mRNA is purified from a cell sample. SLAMP cDNA is preferably generated from such an RNA sample. Quantification of mRNA levels of SLAMP may be useful for the diagnosis of diseases or conditions correlated with abnormal expression of the protein, or to monitor regulation of SLAMP during therapeutic intervention as described herein. Assays for the detection and quantification of mRNA are well known in the art. An example of this method comprises the steps of: isolating RNA from a brain sample of a subject, measuring SLAMP mRNA level by quantitative RT-PCR, and comparing the expression in the subject sample to that of a control sample. Polynucleotide probes or primers used for the detection of SLAMP mRNA by hybridization or PCR amplification may be designed from the polynucleotide of SEQ ID NO:37 by methods well known in the art.

[0669] In another embodiment, the invention relates to methods and compositions for detecting SLAMP and quantifying its level of expression in a biological sample. These methods may be useful for the diagnosis of conditions or diseases characterized by altered or abnormal expression of the protein of the invention, or in assays to monitor subjects being treated with SLAMP, agonists or antagonists. A preferred method comprises contacting an antibody which specifically binds to SLAMP with a brain sample from a mammalian subject, preferably human, and determining the level of

SLAMP in the subject sample compared to a control level representative of a healthy subject, wherein an altered or abnormal level of SLAMP in the subject sample relative to the control level indicates that the subject has the disease or is at an elevated risk of developing the disease. The antibody used can be either monoclonal or polyclonal and can be labeled directly or indirectly for quantification of 5 immune complexes by methods well known to those skilled in the art, for example by ELISA or radioimmunoassays. Preferred specific antibodies directed to SLAMP are generated using the polypeptide fragments listed above. Diagnostic assays to detect the protein of the invention may comprise a biopsy, *in situ* assay of cells. In addition, assays may be conducted upon cellular extracts.

[0670] An embodiment of the invention provides for a method of screening test substances for 10 modulators of SLAMP expression. This method comprises the steps of: i) contacting a cell with a test substance; and ii) comparing SLAMP expression in the cell after exposure to the test substance to that of an unexposed control cell. SLAMP expression is determined by methods common to the art or included herein, by detecting SLAMP polynucleotides or polypeptides. An example of this method comprises the steps of: i) culturing two equivalent cell samples; ii) adding a test substance to one of 15 the cultures and not the other; iii) harvesting both cultures at a specified time; iv) purifying the mRNA from each sample of cells; and v) detecting the level of SLAMP mRNA in each sample. Detection may be accomplished by Northern blot, RT-PCR, or another method common to the art. The invention provides for design and use of specific polynucleotide probes and primers, as discussed herein. An additional example comprises the steps of: i) having two equivalent cultures of cells; ii) adding a test 20 substance to one of the cultures and not the other; iii) harvesting both cultures; iv) purifying the protein from each sample of cells; and v) detecting the level of SLAMP polypeptides in each sample. Detection may be accomplished by enzyme-linked immunoabsorbent assay (ELISA), western blot, radioimmunoassay (RIA), or another method common to the art. The invention provides for design and use of specific antibodies and antibody fragments, as discussed herein. Substances that increase 25 SLAMP expression or activity (agonists or activators) may be used to treat or prevent neurodegenerative diseases affecting the limbic system or abnormal cell proliferation in the non-limbic region. Substances that decrease SLAMP expression or activity (antagonists or inhibitors) may be used to treat or prevent conditions associated with abnormal cell proliferation in the limbic region or neurodegenerative diseases affecting primarily the non limbic system. Methods utilizing SLAMP 30 agonists and antagonists are included herein.

[0671] A preferred embodiment of the invention provides a method of screening for test substances that bind SLAMP polypeptides. This method comprises the steps of: i) contacting a test substance with a SLAMP polypeptide or fragment thereof under conditions that allow binding; and ii)

detecting the binding of the test substance by methods common to the art (e.g., competitive antibody-based methods such as coimmunoprecipitation and Western blotting). Included in this method are test substances that are conjugated to an antibody, antibody fragment, cell-type specific ligand or a portion thereof. This method may be used for purification and detection of SLAMP polypeptides, as well as 5 for the screening methods discussed below.

[0672] A further preferred embodiment of the invention provides a method of screening test substances that bind to SLAMP for agonists of SLAMP activity. This method comprises the steps of: i) contacting a cell with SLAMP and the substance to be tested; and ii) comparing SLAMP biological activity after exposure to the test substance to that of a control exposed only to SLAMP. Measure of 10 SLAMP biological activity may be assessed by monitoring neurite outgrowth. An example of an in vitro assay to measure SLAMP activation comprises the steps of culturing hippocampal neurons in vitro, contacting the cells with SLAMP and the test substance, and measuring neurite number/cell or neurite extension as described by Marg et al. (1999; Journal of Cell Biology, vol.145, pp 865-876, which disclosure is hereby incorporated by reference in its entirety), an increase in the test sample 15 compared to the control sample being indicative of an activating effect of the test substance on SLAMP.

[0673] A further preferred embodiment of the invention provides a method of screening test substances that bind to SLAMP for antagonists of SLAMP activity. This method comprises the steps of: i) contacting a cell with SLAMP and a test substance; and ii) comparing SLAMP biological activity 20 after exposure to the test substance to that of an unexposed control. SLAMP activity may be measured as described above, a decrease in SLAMP activity in the test sample compared to the control sample being indicative of an inhibitory effect of the test substance on SLAMP.

[0674] Another embodiment of the invention is also directed to methods to treat or prevent a disease state of a mammal, preferably a human, associated with excessive neural growth in the limbic 25 region comprising administering to said mammal antagonistic antibodies, ribozymes or antisense vectors or oligonucleotides. Preferably, the disease state to be treated by such methods is brain cancer. A preferred method comprises injecting into an individual antagonistic antibodies directed against SLAMP. Polynucleotide or oligonucleotide sequences complementary to SLAMP polynucleotide sequences may also be used to decrease or inhibit SLAMP expression by limbic neurons. Expression 30 vectors derived from retroviruses, adenovirus, herpes, or vaccinia viruses may be used for the delivery of the polynucleotide sequences complementary to those of the invention to the targeted cell population. Many methods for introducing vectors into cells or tissues are available and equally suitable for use either in vivo or ex vivo. Delivery by transfection and liposome injections may be

achieved using methods which are well known in the art. Alternatively, an antagonist of SLAMP expression or activity isolated as described above can be used, such method comprising the step of administering to a subject an antagonist of SLAMP expression or activity. Preferably, the antagonist is delivered to the limbic neurons of the subject, for example, by conjugating the antagonist to a cell specific targeting moiety (e.g., an antibody fragment). Pharmaceutical compositions comprising SLAMP agonists, antisense vectors or oligonucleotides can be administered alone or in combination with at least one other agent, such as stabilizing compound, which may be administered in any sterile biocompatible carrier. They may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intranasal, enteral or rectal means. The compositions may be administered to the subject alone or in combination with other agents, or drugs.

[0675] It is another object of the invention to provide methods for treating or preventing neuropathologies in a mammal, preferably human, involving the limbic system. This can be achieved by increasing SLAMP expression or activity in vivo using the protein of the invention or fragments thereof or polynucleotides encoding the present protein. A preferred method comprises introducing in vivo into a subject an effective amount of SLAMP polypeptides with a pharmaceutically acceptable carrier. Increased expression of the protein of the invention may also be achieved by introducing into the cells of a subject the sequences encoding the protein of the invention. A preferred method ex-vivo comprising: i) introducing to neural stem cells a polynucleotide sequence encoding SLAMP, and ii) injecting the recombinant cells into an individual. Alternatively, an agonist of SLAMP expression or activity isolated as described above can be used, such method comprising the step of administering to a subject an agonist of SLAMP expression or activity. SLAMP agonists in a physiologically acceptable solution may be delivered by methods common to the art. Preferably, the SLAMP agonist is delivered in vivo to the limbic neurons of the subject. These method are useful for prevention and treatment of neurodegenerative diseases and neuropathies including, but not limited to, Alzheimer's disease, Parkinson's disease, Huntington's disease, cerebral ischaemia, limbic encephalitis, cerebellar ataxia, schizophrenia, and Tourette syndrome.

Protein of SEQ ID NO:40 (Internal designation Clone 495718_160-26-2-0-E12-F)

[0676] The cDNA of Clone 495718_160-26-2-0-E12-F (SEQ ID NO: 39) encodes SAP-MU-
30 10 of SEQ ID NO:40, comprising the amino acid sequence:
MYALFLLASLLGAALAGPVGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKS
LPCDICKDVTAAGDMLKDNAMEEILVYLEKTCDWLPKPMSASCCKEIVDSYLPVILDIIKGE
MSRPGEVCSALNLCESLQKHLAELNHQKQLESNKIPELDMDTEVVAPFMANIPLLLYPQDGPRS

KPQPKDNGDVCQDCIQMVTIDIQTAVRTNSTFVQALVEHVKEECDR LGPGMADICKNYISQYS
EIAIQMMMHMQDQQPKEICALVGF CDEVKEMPMQLVPAKVASKNIPALELVEPIKKHEVP
AKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMC SKLPKSLSEECQE VVDTYGSSILSIL
LEEVSP ELVCSMLHLC SGTRLPALTVHVTQPKDGGFCEVCKKL VGYLDRNLEKNSTKQEILA
5 ALEKGCSFLPD PYQKQCDQFVAEYEPVLIEILVEVWILPSCA. Accordingly, it will be
appreciated that all characteristics and uses of the polypeptides of SEQ ID NO:40 described
throughout the present application also pertain to the polypeptides encoded by the nucleic acids
comprising the human cDNA in Clone 495718_160-26-2-0-E12-F. In addition, it will be appreciated
that all characteristics and uses of the polynucleotides of SEQ ID NOS:39 described throughout the
10 present application also pertain to the nucleic acids comprising the human cDNA in Clone
495718_160-26-2-0-E12-F. A preferred embodiment of the invention is directed toward the
compositions of SEQ ID NO:40, SEQ ID NO:39, and Clone 495718_160-26-2-0-E12-F. Preferred
SAP-MU-10 polypeptide fragment for uses in the methods described below include the Sap-D10
polypeptide comprising the amino sequence of:
15 DGGFCEVCKKL VGYLDRNLEKNSTKQEILA ALEKGCSFLPD PYQKQCDQFVAEYEPVLIEILV
EVWILPSCA. Also preferred are polypeptide fragments comprising the seven C-terminal amino
acids, having a biological activity as described herein and the polynucleotides encoding the fragments.

[0677] The protein of the invention, SAP-MU-10, is a novel isoform of Saposin (GenBank accession number P07602). Three different isoforms of Saposin have been described. The 472 amino 20 terminal amino acids of SAP-MU-10 are identical to the SAP-MU-9 isoform and the 7 carboxyl-terminal amino-acids are unique to SAP-MU-10.

[0678] SAP-MU-10 is a precursor for 4 different proteins, and proteolytic cleavage of SAP-MU-10 gives rise to the Sap-A10, Sap-B10, Sap-C10 and Sap-D10 proteins. The unprocessed SAP-MU-10 precursor binds to sphingolipids and promotes ganglioside transport from liposomes to 25 membrane. The processed proteins are non-enzymatic activators of the hydrolysis of sphingolipids. Sphingolipids are important components of the myelin sheath, a structure that protects and insulates nerve fibers. SAP-MU-10 and Sap-D10 play a major role both in myelination of neurons and in storage of sphingolipids. SAP-MU-10 is expressed and processed in various tissues, being expressed at especially high level in the nervous system. Furthermore, the unprocessed SAP-MU-10 precursor is 30 secreted in the cerebrospinal plasma. The SAP-MU-10 protein binds to sphingolipids and gangliosides, and displays neurotrophic properties. The Sap-D10 protein binds to sphingomyelin and ceramide, and displays a sphingomyelin phosphodiesterase activator activity.

[0679] As used herein, *SAP-MU-10* refers to the polypeptide sequence of SEQ ID NO:40. As used herein, a *SAP-MU-10 polypeptide* refers to a fragment of polypeptide sequence of SEQ ID NO:40 having either sphingolipid or ganglioside binding activity, or neurotrophic properties. As used herein, a *Sap-D10 polypeptide* refers to a polypeptide comprising the Sap-D10 polypeptide described above 5 and having sphingomyelinase phosphodiesterase activator activity.

[0680] An embodiment of the invention is directed to a composition comprising a SAP-MU-10 polypeptide sequence of SEQ ID NO:40.

[0681] A further embodiment of the invention is directed to a composition comprising a SAP-MU-10 polypeptide fragment having biological activity.

10 [0682] A further embodiment of the invention is directed to a composition comprising a Sap-D10 polypeptide fragment having biological activity.

[0683] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence of SEQ ID NO:39 encoding a SAP-MU-10 polypeptide.

15 [0684] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence of SEQ ID NO:39 encoding a Sap-D10 polypeptide.

[0685] A further embodiment of the invention is directed to a composition comprising an antibody that specifically recognizes a SAP-MU-10 polypeptide. Preferably, the antibody recognizes an epitope comprising one or more of the 7 C-terminal amino acids of SAP-MU-10 and Sap-D10, wherein one or more of the seven C-terminal amino acids is required for antibody binding. Preferably, 20 the antibody binds specifically to SAP-MU-10 and to Sap-D10 but not to other saposin isoforms. Alternatively, the antibody binds specifically to SAP-MU-10 but not to Sap-D10 and other saposin isoforms. Alternatively, the antibody binds specifically to Sap-D10 but not to SAP-MU-10 and other saposin isoforms.

[0686] A preferred aspect of the invention is a host cell recombinant for polynucleotides 25 encoding a SAP-MU-10 polypeptide or a biologically active fragment thereof. An additional preferred aspect is a host cell recombinant for polynucleotides capable of directing SAP-MU-10 expression. Preferably, the polynucleotides capable of directing SAP-MU-10 expression are located in the 5' regulatory region of the SAP-MU-10 gene. Further preferably, these polynucleotides are located within 500 base pairs of the SAP-MU-10 coding region. These polynucleotides preferably comprise a 30 promoter sequence. Techniques known in the art for introducing polynucleotide sequences to endogenous sequences are described in U.S. Patent 5641670 and PCT WO9629411, which disclosures

are hereby incorporated by reference in their entireties. SAP-MU-10 protein produced by said host cell may be used for in vitro detection and purification methods as well as diagnosis and in vivo applications.

[0687] Another embodiment relates to a method of producing SAP-MU-10 polypeptides comprising the steps of: i) transfecting a mammalian host cell with a recombinant expression vector comprising a polynucleotide of the present invention, and ii) purifying the produced protein. The purification of the protein can be done following any technique well-known to those skilled in the art. Preferably, an antibody directed against SAP-MU-10 or part thereof may be bound to a chromatographic support to form an affinity chromatography column.

[0688] An embodiment of the invention is a method comprising the step of binding a SAP-MU-10 polypeptide with a sphingolipid or a ganglioside molecule under conditions that allow binding of a SAP-MU-10 polypeptide to said molecule. In a preferred embodiment, a SAP-MU-10 polypeptide is used to purify sphingolipid or ganglioside molecules. In such method, a SAP-MU-10 polypeptide is preferably covalently or non-covalently attached to a solid matrix and allowed to bind sphingolipids and gangliosides using techniques well-known in the art. This method comprises the steps of: i) contacting SPA-MU-10 with a sphingolipid or ganglioside molecule in solution; ii) washing the solid matrix to remove non-binding contaminants; and iii) eluting the sphingolipid or ganglioside using more stringent conditions. Purifying sphingolipids and gangliosides can for example be useful for confirming deficiency of lysosomal enzyme activity affecting ganglioside breakdown, or for monitoring ganglioside catabolism in cultured fibroblasts loaded with ganglioside.

[0689] Additional aspects of this embodiment include methods of using SAP-MU-10 polypeptides to detect and quantify sphingolipids and gangliosides using techniques common in the art. Such methods comprise the steps of: i) obtaining a biological sample suspected of containing sphingolipids or gangliosides; ii) contacting such sample with a SAP-MU-10 polypeptide under conditions allowing binding of SAP-MU-10; and iii) detecting the presence or absence of sphingolipids and gangliosides by detecting SAP-MU-10. Preferably, the SAP-MU-10 polypeptide is covalently attached to a detectable compound (e.g., enzymatic substrates, or fluorescent, luminescent, and radioactive compounds) (E.G....). Alternatively, a detectable SAP-MU-10-specific antibody may be used to detect SAP-MU-10. This embodiment is useful, for example, as a diagnostic tool for quantifying the amount of sphingolipids in cerebrospinal plasma. Such diagnostic tool may be useful to diagnose sphingolipidoses and various lysosomal storage disorders (LSDs) such as, e.g., cystinosis, Gaucher's disease, multiple sulfatase deficiency, Niemann-Pick disease, Pompe's disease and Wolman's disease.

[0690] In another embodiment, the invention is directed to a method of detecting SAP-MU-10 polypeptides in a biological sample, said method comprising the steps of: i) contacting a biological sample with an antibody or antigen binding antibody fragment that specifically binds SAP-MU-10 polypeptide; and ii) detecting the antigen-antibody complex formed. The antibody or antigen binding antibody fragment may be monoclonal or polyclonal. In addition, the antibody or antigen binding antibody fragment may be primarily or secondarily labeled by a detectable compound (e.g., radioactive, fluorescent, luminescent, or enzymatic) common in the art.

[0691] In some embodiments, the invention also concerns a diagnostic kit for detecting *in vitro* the presence of a SAP-MU-10 polypeptide. Such kit comprises: i) a polyclonal or monoclonal antibody or fragment thereof that specifically binds a SAP-MU-10 polypeptide; and optionally, ii) a reagent allowing the detection of the antigen-antibody complexes formed. Preferably, the antibody or antibody fragment is detectably labeled. Such labels include fluorescent, luminescent, and radioactive compounds, as well as enzymatic substrates. The optional reagent may provide a detectable signal and either bind to the antibody or react with the label on such antibody. Preferably, the kit containing antibodies for SAP-MU-10 is used to diagnose LSDs such as those listed herein. Alternatively, the kit of the present invention is used for monitoring the condition of a patient receiving treatment for a LSD. To detect such disorders, an appropriate biological sample (blood or cerebrospinal plasma for example) can be tested to determine the level of SAP-MU-10 being produced. A kit for detecting LSDs comprising an anti-SAP-MU-10 antibody may also comprise other antibodies such as those described in PCT application WO00/55632, which disclosure is incorporated by reference in its entirety. Alternatively, an anti-Sap-D10 antibody may be used in such method.

[0692] An embodiment of the invention provides for a method of screening test substances for modulators of SAP-MU-10 expression. This method comprises the steps of: i) contacting a cell with a test substance; and ii) comparing SAP-MU-10 expression in the cell after exposure to the test substance to that of an unexposed control cell. SAP-MU-10 expression is determined by methods common to the art or included herein, by detecting SAP-MU-10 polynucleotides or polypeptides. An example of this method comprises the steps of: i) culturing two equivalent cell samples; ii) adding a test substance to one of the cultures and not the other; iii) harvesting both cultures at a specified time; iv) purifying the mRNA from each sample of cells; v) comparing the level of SAP-MU-10 mRNA in each sample by Northern blot, RTPCR, or another method common to the art. The invention provides for design and use of specific polynucleotide probes and primers, as discussed herein. An additional example comprises the steps of: i) having two equivalent cultures of cells; ii) adding a test substance to one of the cultures and not the other; iii) harvesting both cultures; iv) purifying the protein from each

sample of cells; v) comparing the level of SAP-MU-10 polypeptides in each sample by Western blot, immunohistochemistry, or another method common to the art. The invention provides for design and use of specific antibodies and antibody fragments, as discussed herein.

[0693] In another embodiment, a SAP-MU-10 polypeptide or a fragment thereof, may be used 5 to screen for compounds that activate or inhibit SAP-MU-10 activity. This method comprises the steps of: i) contacting a SAP-MU-10 polypeptide or fragment thereof with a test substance, and ii) monitoring SAP-MU-10 activity. SAP-MU-10 activity may be monitored by binding assays with gangliosides as described by Hiraiwa *et al.* (Proc Natl Acad Sci U S A. 89:11254-8 (1992)), which disclosure is incorporated by reference in its entirety.

[0694] In another embodiment, a Sap-D10 polypeptide may be used to screen for compounds 10 that activate or inhibit Sap-D10 activity. This method comprises the steps of: i) contacting a Sap-D10 polypeptide or fragment thereof with a test substance, and ii) monitoring Sap-D10 activity. Sap-D10 activity may for example be determined by monitoring sphingomyelinase activity as described by Morimoto *et al.* (Biochem Biophys Res Commun 156:403-10 (1988)), which disclosure is 15 incorporated by reference in its entirety.

[0695] Test substances that decrease SAP-MU-10 expression or activity are defined as inhibitors or antagonists of SAP-MU-10. Test substances that decrease Sap-D10 activity are defined as inhibitors or antagonists of Sap-D10. Test substances that increase SAP-MU-10 expression or 20 activity or that increase Sap-D10 activity are defined as activators or agonists. Agents which modulate the expression or activity of SAP-MU-10 or Sap-D10 activity of the subject invention include, but are not limited to, antisense oligonucleotides, ribozymes, and antibodies. These agents may be made and used according to methods well known in the art.

[0696] One embodiment of the present invention relates to a method for increasing myelination, stimulating neural cell outgrowth, or slowing down a process of demyelination or neural 25 degeneration. The method comprises the step of: contacting a cell with a composition comprising a SAP-MU-10 polypeptide or an agonist thereof. Preferred cells are neurons and oligodendrocytes. The cells can be treated *in vitro* or *in vivo* by directly administering the composition to the cells. In a preferred embodiment, a composition comprising a SAP-MU-10 polypeptide or an agonist thereof is introduced to an individual suffering from a demyelinating disorder (e.g., multiple sclerosis, virus- 30 induced inflammatory demyelination, leukoencephalopathies and leukodystrophies such as Krabbe's disease, metachromatic leukodystrophy, ALD, Canavan disease and Alexander disease). Preferably, the amount of agonist or polypeptide used is effective to increase myelination in an individual. In

another preferred embodiment, a composition comprising a SAP-MU-10 polypeptide or an agonist thereof is introduced to an individual suffering from a neurodegenerative disorder (e.g., retinal neuropathy, Alzheimer's disease, Parkinson's disease, stroke, post-polio syndrome and amyotrophic lateral sclerosis). Preferably, the amount of agonist or polypeptide used is effective to stimulate neural 5 cell outgrowth.

[0697] In another embodiment, a Sap-D10 polypeptide or a Sap-D10 agonist is used to activate sphingomyelinase diphosphoesterase *in vitro* or *in vivo*. A preferred embodiment is directed to a method comprising the step of: contacting a sphingomyelinase diphosphoesterase with a Sap-D10 polypeptide or a Sap-D10 agonist. Another preferred embodiment is directed to a method comprising 10 the step of: contacting a cell with a composition comprising a Sap-D10 polypeptide or an agonist thereof. Still another preferred embodiment relates to a method comprising the step of: administering a composition comprising a Sap-D10 polypeptide or a Sap-D10 agonist in a physiologically acceptable composition to an individual. Preferably, such a method is directed to treat an individual suffering from Tay-Sachs disease associated with Saposin D deficiency. Preferably, the amount of Sap-D10 15 agonist or polypeptide used is effective to increase sphingomyelinase phosphodiesterase activity in the individual suffering from Tay-Sachs disease.

[0698] The composition of the invention may comprise a pharmaceutically acceptable carrier such as those described herein and may be administered to an individual by any technique known in the art. Preferably, the composition is introduced by injection (e.g., direct intracranial injection, 20 injection to the cerebrospinal fluid, local injection to peripheral neural tissue or systemic injection). Even more preferably, the composition is introduced in a patient by one of the methods described in U.S. Patent 5,714,459, and U.S. Patent 4,902,505 which disclosure is incorporated by reference in its entirety.

[0699] Another embodiment is directed to a method of using SAP-MU-10 polynucleotides 25 comprising the steps of: i) constructing a recombinant molecule comprising a nucleic acid sequence encoding a SAP-MU-10 polypeptide that allows expression of SAP-MU-10 or fragment thereof under suitable physiological conditions, and ii) introducing this recombinant molecules into a cell, a mammal or a human. Preferably, the recombinant molecule is introduced into a neural cell.

[0700] Recombinant molecules comprising a nucleic acid sequence encoding a SAP-MU-10 30 polypeptide may be directly introduced into cells or tissues *in vitro* using delivery vehicles such as retroviral vectors, adenoviral vectors and DNA virus vectors. They may also be introduced into cells *in vivo* using physiological techniques such as microinjection and electroporation or chemical methods

such as coprecipitation and incorporation of DNA into liposomes. Recombinant molecules may also be delivered in the form of an aerosol or by lavage. The SAP-MU-10 polynucleotides may also be applied intracellularly such as by direct injection into cells. Methods for carrying out SAP-MU-10 transfer into a cell and grafting in the brain are described in U.S. Patent 5,762,926, which disclosure is hereby incorporated by reference in its entirety. Such a method can be used to treat, e.g., multiple sclerosis, leukoencephalopathies and leukodystrophies.

[0701] Another embodiment of the invention relates to transgenic animals and methods of using a SAP-MU-10 polynucleotide sequence or part thereof to establish transgenic animals (*D. melanogaster*, *M. Musculus*), by any method familiar to those skilled in the art. By modulating *in vivo* the expression of the transgene with drugs or modifier genes (activator or suppressor genes), animal models can be developed that mimic human disorders such as LSDs. These animal models would thus allow the identification of potential therapeutic agents for treatment of the disorders. In addition, recombinant cell lines derived from these transgenic animals may be used for similar approaches *ex vivo*.

15 Protein of SEQ ID NO:42 (Internal designation Clone 612386_187-9-4-0-B2-F)

[0702] The cDNA of Clone 612386_187-9-4-0-B2-F (SEQ ID NO: 41) encodes cytogram of SEQ ID NO:42, comprising the amino acid sequence: MELCRSLALLGGSLGLMFCLIALSTDFWFEAVGPTHSAHSGLWPTGHGDIISGHGPLVSTTAA FAAGKDSGLDWGIASQRIPAEELSHLSCPCPQPSPWWPWRCRTPASGGTSLHTPRSRPSSPGP 20 STWAGSQLSSCSVQVP. Accordingly, it will be appreciated that all characteristics and uses of the polypeptides of SEQ ID NO:42 described throughout the present application also pertain to the polypeptides encoded by the nucleic acids comprising the human cDNA in Clone 612386_187-9-4-0-B2-F. In addition, it will be appreciated that all characteristics and uses of the polynucleotides of SEQ ID NOs:41 described throughout the present application also pertain to the nucleic acids included in 25 Clone 612386_187-9-4-0-B2-F. A preferred embodiment of the invention is directed toward the compositions of SEQ ID NO:42, SEQ ID NO:41, and Clone 612386_187-9-4-0-B2-F. Preferred cytogram polypeptides for uses in the methods described below include the polypeptides comprising the amino sequence of: AHSGLWPTGHGDIISGHGPLVSTTAAFAAGKDSGLDWGIASQRIPAEELSHLSCPCPQPSPWW 30 WPWRCTPASGGTSLHTPRSRPSSPGPSTWAGSQLSSCSVQVP. Also preferred are polypeptide fragments having a biological activity as described herein and the polynucleotides encoding the fragments.

[0703] The protein of the invention, cytogram, is a splice variant of GMP-17 (or NKG7, GenBank accession number Q16617). Cytogram cDNA lacks the second exon of GMP-17 cDNA. Cytogram is a 142 amino-acid long protein, and GMP-17 is 165 amino acid long. The 53 amino-terminal amino acids are identical between the two proteins, and the 89 carboxyl-terminal amino acids 5 are unique to cytogram. Cytogram displays a signal peptide (MELCRSLALLGGSLGLMFCLIALSTDFWFEAVGPTHS), a transmembrane domain (GDIISGHGPLVSTTAFAAGK), and a PEC family methallothionein domain (SCPCPQPSPWWWPWRCTPASGGTSLHTPRSRPSSPGPSTWAGSQLSSCSV) that binds to zinc ions.

[0704] Cytogram is a cytotoxic granule membrane protein that is constitutively and specifically expressed in NK cells (NKs) and cytotoxic T lymphocytes (CTLs). NK and CTL degranulation results in translocation of cytogram from granules to the plasma membrane. Cytogram contributes to the formation of junctions between effector cells and target cells following exocytosis. Furthermore, once located on the plasma membrane, cytogram regulates ion channels required for 15 cytotoxicity of NKs and CTLs and promotes cytolytic effector function.

[0705] An embodiment of the invention is directed to a composition comprising a cytogram polypeptide sequence of SEQ ID NO:42.

[0706] A further embodiment of the invention is directed to a composition comprising a cytogram polypeptide fragment having biological activity of promoting NK and CTL cytotoxicity.

[0707] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence of SEQ ID NO:41 encoding a cytogram polypeptide.

[0708] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence encoding a cytogram polypeptide fragment having biological activity.

[0709] A further embodiment of the invention is directed to a composition comprising an 25 antibody recognizing a cytogram polypeptide sequence of SEQ ID NO:42 or a cytogram polypeptide fragment having biological activity. Preferably, the antibody recognizes a non-linear epitope or an epitope located within the 89 C-terminal amino acids of cytogram. Preferably, the antibody binds to cytogram but not to GMP-17. Preferably, the antibody recognizes the AGKDSGLD amino acid sequence.

[0710] A further embodiment of the present invention relates to a method of binding activated 30 NKs and CTLs comprising the step of: contacting activated NKs and CTLs with an antibody that

specifically binds to cytogram polypeptides under conditions that allow binding of cytogram polypeptides to the antibody. In a preferred embodiment, antibodies that specifically bind to cytogram polypeptides can be used to purify activated NKs and CTls. In such method, the antibody is preferably covalently or not covalently attached to a solid matrix and allowed to bind cytogram polypeptides 5 using techniques well known in the art. This method comprises the steps of: i) contacting activated NKs and CTls with an antibody that specifically binds to cytogram polypeptides under conditions that allow binding of cytogram polypeptides to the antibody, ii) washing the solid matrix to get rid of contaminants, and iii) eluting the cells of interest using more stringent conditions. Such a method can for example be useful for purifying NKs and CTls from cancer patients prior to *in vitro* expansion.

10 [0711] A preferred aspect of the invention is a host cell recombinant for polynucleotides encoding a cytogram polypeptide or a biologically active fragment thereof. An additional preferred aspect is a host cell recombinant for polynucleotides capable of directing cytogram expression. Preferably, the polynucleotides capable of directing cytogram expression are located in the 5' regulatory region of the cytogram gene. Further preferably, these polynucleotides are located within 15 500 base pairs of the cytogram coding region. These polynucleotides preferably comprise a promoter sequence. Techniques known in the art for introducing polynucleotide sequences to endogenous sequences are described in U.S. Patent 5641670 and PCT WO9629411, which disclosures are hereby incorporated by reference in their entireties. Cytogram protein produced by said host cell may be used for in vitro detection and purification methods as well as diagnosis and in vivo applications.

20 [0712] One embodiment of the present invention is directed to methods of detecting or quantifying activated NKs and CTls comprising the steps of: i) contacting a body fluid, a tissue sample or a mammalian cell culture with an antibody that specifically binds to cytogram polypeptides, and ii) detecting the antibody in the sample using any detectable signal. Compounds such as, e.g., alkaline phosphatase, peroxidase, FITC, rhodamine, Texas-Red, biotin and digoxigenin can be used to 25 provide a detectable signal. The detection of the signal may be carried out using immunohistological or immunofluorescing processes that are well-known to those skilled in the art. The antibody that specifically binds to cytogram polypeptides may be directly labeled with the compound giving the detectable signal. Alternatively, the antibody that specifically binds to cytogram polypeptides is not labeled and detection of the signal occurs indirectly by using labeled secondary antibodies. In a 30 preferred embodiment, detecting activated NKs and CTls can be used to measure the effect of a test compound on CTL or NK activity in mammalian cell cultures. In another preferred embodiment, such methods can be used to monitor the effects of a treatment aiming to increase or decrease CTL or NK activity in a patient, or to detect the beginning of a graft rejection reaction in a patient.

[0713] An embodiment of the present invention relates to methods of using cytogram or fragment thereof as a marker of cytotoxic granules comprising the steps of: i) contacting a body fluid, a tissue sample or a mammalian cell culture with an antibody that specifically binds to cytogram polypeptides, and ii) detecting the localization of the antibody in the sample using any detectable signal. Compounds such as, e.g., FITC, rhodamine, Texas-Red, [³⁵S]methionine, [³⁵S]cysteine and bisbenzimidazole can be used to provide a detectable signal. The detection of the signal may be carried out using immuno-electron microscopy or fluorescence microscopy. The antibody that specifically binds to cytogram polypeptides may be directly labeled with the compound giving the detectable signal. Alternatively, the antibody that specifically binds to cytogram polypeptides is not labeled and 10 detection of the signal occurs indirectly by using labeled secondary antibodies. Such methods of using cytogram or part thereof as a marker of cytotoxic granules include those described by Medley *et al.* (Proc Natl Acad Sci U S A. 93:685-9 (1996)), which disclosure is incorporated by reference in its entirety. CTLs and NKs form the major defense of higher organisms against virus-infected and transformed cells and a key function of CTLs is to detect and eliminate potentially harmful cells by 15 lysis. Thus, a marker for detecting degranulation of NKs and CTLs is very useful when screening for drugs to treat or monitor the status or progression of, for example, infection with intracellular pathogens, graft versus host disease, susceptibility to transplantable and spontaneous malignancies, lymphoid homeostasis, the tendency to auto-immune diseases or when studying these diseases.

[0714] Another embodiment relates to a method of producing cytogram polypeptides 20 comprising the steps of: i) obtaining a cell capable of expressing a cytogram polypeptide; ii) growing said cell under conditions suitable to produce said polypeptide; and iii) purifying said polypeptide.. The purification of the protein can be done following any technique well-known to those skilled in the art. Preferably, an antibody directed against cytogram or part thereof may be bound to a chromatographic support to form an affinity chromatography column. Even more preferably, the 25 antibody binds to cytogram but not to GMP-17. The cell capable of expressing a cytogram polypeptide may be obtained by any of the techniques well-known to those skilled in the art. A host cell may be transfected with a recombinant expression vector comprising a polynucleotide of the present invention. Alternatively, an heterologous promotor may be used. Preferably, the host cell is a mammalian host cell.

[0715] Another preferred embodiment relates to a method of screening test substances for 30 modulators of cytogram expression. This method comprises the steps of: i) contacting a cell with a test substance; and ii) comparing cytogram expression in the cell after exposure to the test substance to that of an unexposed control cell. Cytogram expression is determined by methods common to the art

or included herein. Methods of determining cytogram expression include but are not limited to methods of quantifying cytogram polynucleotides (e.g., detection of cytogram mRNA by northern blot or RTPCR) or to methods of quantifying cytogram polynucleotides (e.g., detection of cytogram proteins by western blot or immunochemistry). Preferably, the test substance modifies the expression 5 of cytogram in a specific cell type while not in others. Most preferably, the test substance modifies cytogram expression specifically in NKs and CTLs.

[0716] A further embodiment of the present invention is directed to test substances for modulators of cytogram activity and to a method of screening for such test substances comprising the steps of: i) contacting a cell with a test substance, ii) determining cytogram activity, and iii) comparing 10 cytogram activity in the cell after exposure to that of an unexposed control cell. Cytogram activity can for example be monitored by studying cytotoxic activity of NKs using the cytotoxic assay described in U.S. Patent 5,229,494, which disclosure is hereby incorporated by reference in its entirety. Preferably, cytogram activity is studied in NKs or CTLs. Most preferably, a substance that modulates cytogram activity is administered to an individual to modulate NK or CTL cytotoxicity.

15 [0717] Test substances that decrease cytogram expression or activity are defined as inhibitors or antagonists of cytogram. Test substances that increase cytogram expression or activity are defined as activators or agonists of cytogram. Tests substances that modulate the expression or activity of cytogram include but are not limited to chemical compounds, oligonucleotides, ribozymes and antibodies. These substances may be made and used according to methods well known in the art.

20 [0718] An embodiment of the present invention relates to a method of using the compositions of the present invention to prevent or reduce in severity disorders caused as a result of NK or CTL cytotoxicity. In a preferred embodiment, such a method comprises the step of administering to an individual a cytogram antagonist in a physiologically acceptable composition. Preferably, such methods for reducing NK and CTL activation are directed to treat allergy and asthma, and to prevent 25 on-going immune responses. More particularly, such methods can be used to prevent or reduce in severity graft rejections, graft versus host diseases and autoimmune diseases (e.g., rheumatoid arthritis, systemic lupus, multiple sclerosis, insulin-dependent diabetes, hepatitis, rheumatoid arthritis, Graves disease), and to induce tolerance to graft transplantation (e.g., transplantation of cells, bone marrow, tissue, solid-organ, bone). A preferred cytogram antagonist for use in such method is an antibody that 30 specifically binds to cytogram polypeptides.

[0719] In another preferred embodiment, an anti-cytogram antibody may be used as a targeting and delivery mechanism for bringing a pharmaceutical agent to activated NKs and CTLs.

Such a method comprises the steps of: i) conjugating an anti-cytogram antibody to a pharmaceutical agent; and ii) administering a composition comprising said conjugate and a pharmaceutically acceptable carrier to an individual suffering from a disorder caused as a result of NK or CTL cytotoxicity. Preferably, an anti-cytogram antibody is used in an antibody-directed enzyme-prodrug therapy comprising the steps of: i) conjugating an anti-cytogram antibody to an enzyme converting a relatively non-toxic compound into a substantially more toxic compound; and ii) administering said relatively non-toxic compound to a patient after a delay that allowed residual enzyme-antibody conjugate to be cleared from the blood. Such a method may be performed as described by Melton *et al.* (*J. Natl. Cancer Inst.* 1996 88:153-65 (1996)), which disclosure is hereby incorporated by reference in its entirety.

[0720] An embodiment of the present invention relates to methods of using the compositions of the present invention to enhance NK or CTL cytotoxicity. A preferred method directed to enhance NK or CTL cytotoxicity comprises the step of introducing a cytogram polypeptide or a cytogram agonist in a physiologically acceptable carrier in an individual suffering from diseases and disorders where a boost to the immune system is desirable. More particularly, such treatments can be used in conjunction with a radiation therapy or a chemotherapy when treating a cancer, and can be used to treat neoplastic disorders (e.g., multiple myeloma, colon cancer, hepatoma) viral infections (e.g., HIV, HBV, HCV, hepatitis, measles and herpes viruses infections), and various immune deficiencies. These immune deficiencies may be genetic (e. g. severe combined immunodeficiency (SCID)) or be caused by various bacterial or fungal infections (e.g. infections by mycobacteria, *Leishmania* spp., malaria spp. and candidiasis).

[0721] Physiologically acceptable carriers can be prepared by any method known by those skilled in the art. Physiologically acceptable carriers include but are not limited to those described in Remington's Pharmaceutical Sciences (Mack Publishing Company, Easton, USA 1985), which disclosure is hereby incorporated by reference in its entirety. Pharmaceutical compositions can be for, e.g., intravenous, topical, rectal, local, inhalant or subcutaneous, intradermal, intramuscular, oral and intracerebral use. The compositions can be in liquid (e.g., solutions, suspensions), solid (e.g., pills, tablets, suppositories) or semisolid (e.g., creams, gels) form. Dosages to be administered depend on individual needs, on the desired effect and the chosen route of administration.

[0722] Another embodiment is directed to a method of using cytogram polynucleotides comprising the steps of: i) constructing a recombinant molecule comprising a nucleic acid sequence encoding a cytogram polypeptide that allows expression of cytogram or part thereof under given physiological conditions, and ii) introducing this recombinant molecules into a cell, a mammal or a

human. Such a method can be used to treat, e.g., the genetic immune deficiencies and neoplastic diseases listed above. Preferably, the recombinant molecule is introduced into NKs or CTLs.

[0723] Still another embodiment is directed to a method of using antisense cytogram polynucleotides comprising the steps of: i) constructing a recombinant molecule comprising a nucleic acid sequence encoding a cDNA that is complementary to cytogram polynucleotides, and ii) introducing this recombinant molecules into a cell, a mammal or a human. Such a method can be used to treat autoimmune diseases. Alternatively, such a method can be used to prevent the autoimmune disorders listed above, graft versus host disease or graft rejection when transplanting an organ. The recombinant molecule may be introduced into NKs, CTLs, or graft cells.

10 [0724] Recombinant molecules comprising a nucleic acid sequence encoding a cytogram polypeptide or a nucleic acid sequence that encodes a cDNA that is complementary to cytogram polynucleotides may be directly introduced into cells or tissues *in vivo* using delivery vehicles such as retroviral vectors, adenoviral vectors and DNA virus vectors. They may also be introduced into cells *in vivo* using physiological techniques such as microinjection and electroporation or chemical methods 15 such as coprecipitation and incorporation of DNA into liposomes. Recombinant molecules may also be delivered in the form of an aerosol or by lavage. The cytogram polynucleotides may also be applied extracellularly such as by direct injection into cells. Preferably, cytogram polynucleotides are introduced into NKs or CTLs.

Protein of SEQ ID NO:44 (Internal designation Clone 1000838982_220-20-4-0-C2-F)

20 [0725] The cDNA of Clone 1000838982_220-20-4-0-C2-F (SEQ ID NO: 43) encodes Tetranab of SEQ ID NO:44, comprising the amino acid sequence: MSAQESCLSLIKYFLVFNLFFFVLGSLIFCFGIWILIDKTSFVSFVGLAFVPLQIWSKVL AISGIF TMGIALLGCVGALKELRCLLGLYFGMLLLLFATQITLGILISTQRASWSEACGTS. Accordingly, it will be appreciated that all characteristics and uses of the polypeptides of SEQ ID NO:44 described 25 throughout the present application also pertain to the polypeptides encoded by the nucleic acids comprising the human cDNA in Clone 1000838982_220-20-4-0-C2-F. In addition, it will be appreciated that all characteristics and uses of the polynucleotides of SEQ ID NOs:43 described throughout the present application also pertain to the nucleic acids included in Clone 1000838982_220-20-4-0-C2-F. A preferred embodiment of the invention is directed toward the 30 compositions of SEQ ID NO:44, SEQ ID NO:43, and Clone 1000838982_220-20-4-0-C2-F. Also preferred are polypeptide fragments having a biological activity as described herein and the polynucleotides encoding the fragments.

[0726] The protein of the invention, Tetranab, is a variant of CD37 (GenBank accession number P11049). CD37 is a molecular facilitator that brings together molecular complexes necessary for T-cell dependant B-cell response. CD37 stabilizes molecular interactions, resulting in a more efficient response. Tetranab is a divergent member of the tetraspanin family, which displays one signal anchor (FNLFFFVLGSLIFCFGIWILI) and two transmembrane domains (VLAISGIFTMGIALLGCVGAL and LYFGMLLLLFATQITLGILIS). Tetranab is specifically expressed on mature B cells. It is not expressed during differentiation of B cells and its expression is down-regulated with activation of B cells. Tetranab acts as a dominant negative inhibitor of CD37-facilitated assembly of functional complexes at B cells surfaces. Thus Tetranab prevents CD37-dependent activation of B cells.

[0727] An embodiment of the invention is directed to a composition comprising a Tetranab polypeptide sequence of SEQ ID NO:44.

[0728] A further embodiment of the invention is directed to a composition comprising a Tetranab polypeptide fragment having biological activity of preventing T-cell dependant B-cell response.

[0729] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence of SEQ ID NO:43 encoding a Tetranab polypeptide.

[0730] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence encoding a Tetranab polypeptide fragment having biological activity of preventing T-cell dependent B-cell response.

[0731] A further embodiment of the invention is directed to a composition comprising an antibody recognizing a Tetranab polypeptide. Preferably, the antibody recognizes an epitope comprising one or more of the 9 C-terminal amino acids of Tetranab, wherein one or more of the 9 C-terminal amino acids is required for antibody binding. Preferably, the antibody binds specifically to Tetranab but not to CD37. As used herein, an anti-Tetranab antibody refers to an Tetranab-specific antibody or an antigen binding fragment thereof.

[0732] The present invention also relates to a method of binding such an antibody to a Tetranab polypeptide comprising the step of: contacting a Tetranab polypeptide with said antibody or an antigen-binding fragment thereof under conditions that allow binding to take place. Such conditions are well known to those skilled in the art. Such methods are useful for detecting Tetranab polypeptides or for purifying B cells as further described herein.

[0733] An embodiment of the present invention is directed to a method of detecting Tetranab polypeptides in a biological sample, said method comprising the steps of: i) contacting a biological sample with an anti-Tetranab antibody; and ii) detecting the antigen-antibody complex formed. The antibody or antibody fragment may be monoclonal or polyclonal. In addition, the antibody or 5 antibody fragment may be primarily or secondarily labeled by any detectable compound (e.g., radioactive, fluorescent, luminescent, or enzymatic) common in the art. Further preferred is a method of using said anti-Tetranab antibodies or antigen binding fragments in a flow cytometric analysis of Tetranab expression by B cells. Such a method can be used as a diagnostic tool for detecting the presence of a Tetranab polypeptide as further described herein.

[0734] Another embodiment of the present the invention is directed to a diagnostic kit for detecting *in vitro* the presence of a Tetranab polypeptide. Such kit comprises: i) a polyclonal or monoclonal antibody or fragment thereof that specifically binds a Tetranab polypeptide; and optionally, ii) a reagent allowing the detection of the antigen-antibody complexes formed. Preferably, the antibody or antibody fragment is detectably labeled as described above. The optional reagent may 15 provide a detectable signal, and either bind to the antibody, or react with the label on said antibody. Preferably, the kit comprising anti-Tetranab antibodies or antigen binding fragments is used for diagnosing non-Hodgkin's lymphomas, or for monitoring the condition of a patient receiving treatment for a non-Hodgkin's lymphoma. Optionally, said kit for diagnosing non-Hodgkin's lymphomas comprises a sample of an individual suffering from lymphoma, as well as a control sample 20 from a normal individual.

[0735] Another preferred embodiment relates to a method for detecting *in vivo* the presence of a Tetranab polypeptide using anti-Tetranab antibodies. Such a method is very useful for imaging of tumor involved sites, and especially for imaging of non-Hodgkin's lymphomas. Preferably, imaging is carried out using ^{111}In or ^{131}I conjugated to an anti-Tetranab antibody as described by Bunn *et al* 25 (Lancet 2:1219-21 (1894)), which disclosure is incorporated by reference in its entirety. This embodiment also relates to a kit for use in tumor imaging such as the kit described above.

[0736] Still another embodiment of the present invention relates to a method using an anti-Tetranab antibody for immunotherapeutic purposes comprising the step of administering a therapeutic dose of a radioactively labeled anti-Tetranab antibody in a pharmaceutically acceptable carrier. Such a 30 method can be applied to treat non-Hodgkin's lymphomas and post-transplant lymphoproliferative disorders. In such a method, labeling with ^{90}Y or ^{131}I is preferred. Radioactive isotopes can be attached to the anti-Tetranab antibody by a number of methods well-known to those skilled in the art, e.g., by covalent attachment of a iodine isotope directly to the antibody or by covalent attachment of a

chelating moiety to the antibody allowing the chelator to coordinate the metal isotope. Routes of administration may vary widely. Preferred modes of administration are intravenous injection and intralymphatic routes of administration such as subcutaneous and intramuscular injection, or catheterization of lymphatic vessels. The radiometric dosage to be applied can vary substantially.

- 5 Methods of administration and dosage regimen for treating lymphomas can be determined as described in US Patent 5,595,721, which disclosure is hereby incorporated by reference in its entirety.

[0737] A preferred aspect of the invention is a host cell recombinant for polynucleotides encoding a Tetranab polypeptide or a biologically active fragment thereof. An additional preferred aspect is a host cell recombinant for polynucleotides capable of directing Tetranab expression. 10 Preferably, the polynucleotides capable of directing Tetranab expression are located in the 5' regulatory region of the Tetranab gene. Further preferably, these polynucleotides are located within 500 base pairs of the Tetranab coding region. These polynucleotides preferably comprise a promoter sequence. Techniques known in the art for introducing polynucleotide sequences to endogenous sequences are described in U.S. Patent 5641670 and PCT WO9629411, which disclosures are hereby 15 incorporated by reference in their entireties. Tetranab protein produced by said host cell may be used for in vitro detection and purification methods as well as diagnosis and in vivo applications.

[0738] An embodiment of the invention provides for a method of screening test substances for modulators of Tetranab expression. This method comprises the steps of: i) contacting a cell with a test substance; and ii) comparing Tetranab expression in the cell after exposure to the test substance to that 20 of an unexposed control cell. Tetranab expression is determined by methods common to the art or included herein, by detecting Tetranab polynucleotides or polypeptides. An example of this method comprises the steps of: i) culturing two equivalent cell samples; ii) adding a test substance to one of the cultures and not the other; iii) harvesting both cultures at a specified time; iv) purifying the mRNA from each sample of cells; v) comparing the level of Tetranab mRNA in each sample by Northern blot, 25 RTPCR, or another method common to the art. An additional example comprises the steps of: i) having two equivalent cultures of cells; ii) adding a test substance to one of the cultures and not the other; iii) harvesting both cultures; iv) purifying the protein from each sample of cells; v) comparing the level of Tetranab polypeptides in each sample by Western blot, immunohistochemistry, or another method common to the art. Preferably, Tetranab expression is measured in B cells.

30 [0739] In another embodiment, a Tetranab polypeptide or a fragment thereof may be used to screen for compounds that modulate Tetranab activity. This method comprises the steps of: i) contacting a Tetranab polypeptide or fragment thereof with a test substance, and ii) monitoring Tetranab activity. Tetranab activity may be determined by monitoring B cell activation. For example,

B cell function can be determined by measuring calcium ion flux across the membrane as described in US Patent 4,987,084, which disclosure is incorporated by reference in its entirety.

[0740] Test substances that decrease Tetranab expression or activity are defined as inhibitors or antagonists of Tetranab. Test substances that increase Tetranab expression or activity are defined as 5 activators or agonists. Agents that modulate the expression or activity of Tetranab activity of the subject invention include, but are not limited to antisense oligonucleotides, ribozymes, and antibodies. These agents may be made and used according to methods well known in the art.

[0741] A further embodiment of the present invention relates to a method of antagonizing CD37 action, thus preventing T-cell dependant B-cell activation. Such a method comprises the step 10 of contacting a B cell with a composition comprising a Tetranab agonist. The B cells can be treated *in vitro* or *in vivo* by directly administering the composition of the present invention to the cells. In a preferred embodiment, this invention provides a method of inhibiting B cell activation in an individual comprising the step of administering to the individual an effective inhibiting amount of a pharmaceutical composition comprising a Tetranab agonist and a pharmaceutically acceptable carrier. 15 A method of inhibiting B cell activation is a useful method for inhibiting the immune response of an individual. Preferably, inhibiting the immune response using a method of the present invention is directed to inhibit a graft rejection reaction, to treat an individual suffering from autoimmune disease (e.g., rheumatoid arthritis, Myasthenia gravis, systemic lupus erythematosus, Graves' disease, idiopathic thrombocytopenic purpura, hemolytic anemia, diabetes mellitus and drug-induced 20 autoimmune diseases such as drug-induced lupus), or to reduce allergic responses (e.g., asthma, hay fever and allergy to penicillin).

[0742] Another embodiment of the present invention relates to a method of favoring CD37 action, thus stabilizing molecular complexes that are necessary for T cell dependant B cell activation. Such a method comprises the step of contacting a B cell with a composition comprising a Tetranab 25 antagonist. In this aspect of the invention, an effective inhibiting amount of a pharmaceutical composition comprising a Tetranab antagonist and a pharmaceutically acceptable carrier may be administered to an individual in order to enhance B cell activation. Such a method is useful for treating disorders where a boost of the immune system is desirable such as immune deficiencies (e.g., severe combined immunodeficiency), various viral, fungal or bacterial infections (e.g., HBV, HCV, 30 HIV, hepatitis, measles and herpes virus infections, mycobacteria, Leishmania spp., malaria spp. and candidiasis infections), and in association with a radiation or chemotherapy.

[0743] Compositions comprising a Tetranab agonist or a Tetranab antagonist may be formulated using a variety of acceptable excipients known in the art. Typically, the compositions are administered topically, orally or by injection, either intravenously or intraperitoneally. Methods to accomplish this administration and to determine the dosage regimen are known to those of ordinary skill in the art.

[0744] Another embodiment relates to a method of producing Tetranab polypeptides comprising the steps of: i) transfecting a host cell with a recombinant expression vector comprising a polynucleotide of the present invention, and ii) purifying the produced protein. The purification of the protein can be done following any technique well-known to those skilled in the art. Preferably, an antibody directed against Tetranab or part thereof may be bound to a chromatographic support to form an affinity chromatography column. Such a purified Tetranab polypeptide may for example be useful for preparing the pharmaceutical compositions described above.

[0745] A preferred aspect of the invention is a host cell recombinant for polynucleotides encoding a Tetranab polypeptide or a biologically active fragment thereof. An additional preferred aspect is a host cell recombinant for polynucleotides capable of directing Tetranab expression. Preferably, the polynucleotides capable of directing Tetranab expression are located in the 5' regulatory region of the Tetranab gene. Further preferably, these polynucleotides are located within 500 base pairs of the Tetranab coding region. These polynucleotides preferably comprise a promoter sequence. Techniques known in the art for introducing polynucleotide sequences to endogenous sequences are described in U.S. Patent 5641670 and PCT WO9629411, which disclosures are hereby incorporated by reference in their entireties. In preferred embodiments, the recombinant molecule is introduced into an *Escherichia coli* cell or into a human B cell. Alternatively, the recombinant molecule is a vector comprising an heterologous promotor that increases Tetranab expression that is introduced into a human cell. In a preferred embodiment, said recombinant molecule is introduced into animals, more particularly mammals. Said recombinant molecule may be introduced into animals in several ways. A preferred method of introducing the recombinant molecule into animals comprises the steps of: i) introducing the recombinant molecule into a host cell by means known in the art such as, e.g., transformation, electroporation, lipofection, microinjection, and transduction, including the use of retroviral vectors, adenoviral vectors and DNA virus vectors; and ii) administering the recombinant cells to an animal. For example, the cells can be administered by infusion. Such a method may be directed to treat, e.g., autoimmune diseases. Alternatively, the recombinant molecule may encode a cDNA that is complementary to SEQ ID NO:43 or part thereof, and the method may be used to treat, e.g., immune deficiencies.

[0746] In another preferred embodiment, said recombinant molecule is used to establish transgenic model animals by any method familiar to those skilled in the art. For example, transgenic mice may be established as described by Hogan *et al.* ("Manipulating the Mouse Embryo: A Laboratory Manual", Cold Spring Harbor Laboratory), which disclosure is incorporated herein by reference. By modulating *in vivo* the expression of the transgene with drugs or modifier genes (activator or suppressor genes), animal models can be developed that mimic human disorders such as those listed above. These animal models thus allow the identification of potential therapeutic agents for treatment of the disorders. In addition, recombinant cell lines derived from these transgenic animals may be used for similar approaches *ex vivo*.

10 Protein of SEQ ID NO:46 (internal designation Clone 500720840_205-20-1-0-B7-F)

[0747] The cDNA of Clone 500720840_205-20-1-0-B7-F (SEQ ID NO:45) encodes PDI protein of SEQ ID NO:46 comprising the amino acid sequence:

MRLRRLALFPGVALLAARLAAASDVLELTDDNFESRISDTGSAGLMLVEFFAPWCGHCKR
LAPEYEAAATRLKGIVPLAKVDCTANTNTCNKYGVSGYPTLKIFRDGEEAGAYDGPRADGI
15 VSHLKKQAGPASVPLRTTEEFKKFISDKDASIVGFFDDSFSEAHSEFLKAASNLRDNYRFAHTN
VESLVNEYDDNGEIILFRPSHLTNKFEDKTVAYTEQKMTSGKIKKFIQENIFGICPHMTEDNK
DLIQGKDLLIAYYDVEDYEKNAKGSNYRRNRVMMVAKKFLDAGHKLNAVASRKTFSHELSD
FGLESTAGEIPVVAIRTAKGEKFVMQEFSRDGKALERFLQDYFDGNLKRYLKSEPIPESNDGP
VKVVVAENFDEIVNNENKDVLIEFYAPWCGHCKNLEPKYKELGEKLSKDPNIVIAKMDATAN
20 DVPSPYEVRGFTIYFSPANKKLNPKKYEGGRELSDFISYQLREATIPPVIQEEKPKKKKAQE
DL. Accordingly, it will be appreciated that all characteristics and uses of the polypeptides of SEQ ID NO:46 described throughout the present application also pertain to the polypeptides encoded by the nucleic acids included in Clone 500720840_205-20-1-0-B7-F. In addition, it will be appreciated that all characteristics and uses of the polynucleotides of SEQ ID NO:45 described throughout the present 25 application also pertain to the nucleic acids comprising the human cDNA in Clone 500720840_205-20-1-0-B7-F. A preferred embodiment of the invention is directed toward the compositions of SEQ ID NO:45, SEQ ID NO:46 and Clone 500720840_205-20-1-0-B7-F. Also preferred are polypeptide fragments having a biological activity as described herein and the polynucleotides encoding the fragments.

30 [0748] The protein of SEQ ID NO:46, PDI, is a polymorphism variant of protein disulfide isomerase A3 precursor (accession numbers U42068). The protein of the invention displays two thioredoxin domain:

ASDVLELTDDNFESRISDTGSAGLMLVEFFAPWCGHCKRLAPEYEAAATRLKGIVPLAKVDC
TANTNTCNKYGVSGYPTLKIFRDGEAGAYDGPRTADGIVSHLKKQAG; and
DGPVKVVVAENFDEIVNNENKDVLIEFYAPWCGHCKNLEPKYKELGEKLSKDPNIVIAKMDA
TANDVPSPYEVRGFPTIYFSPANKKLNPKKYEGGRELSDFISYQLQREAT. Also, the 505 amino
5 acid protein of PDI displays one membrane-spanning segment: SDTGSAGLMLVEFFAPWCGHC.

[0749] Thioredoxin family active site proteins are a superfamily of proteins that participate in redox reactions and are distributed among a wide range of living organisms. The reduced form of thioredoxin is known to activate some enzymes by reducing disulfide bridges that control their activity. In addition, thioredoxin is an electron donor in the reactions sequence that reduces 10 ribonucleotides to deoxyribonucleotides catalyzed by ribonucleotide reductase. It has been reported that in humans, thioredoxin and the cellular redox state modified by thioredoxin play a crucial role in arterial neointima formation in atherosclerosis. Thioredoxin is involved in cellular defense mechanisms against oxidative damage. Thioredoxin has also been implicated in regulating glucocorticoid responsiveness to cellular oxidative stress response pathways. In particular, thioredoxin 15 is capable of sensing and transmitting the redox state of the cell to the glucocorticoid receptor by targeting both the ligand- and DNA-binding domains of the receptor. Human thioredoxin has been suggested to act as a free radical scavenger and has been shown to limit the extent of ischemia reperfusion injury. Multiple in vitro substrates for thioredoxin have been identified, including ribonuclease, choriogonadotropins, coagulation factors, glucocorticoid receptor, and insulin.

[0750] PDI belongs to the protein disulfide isomerase class of thioredoxin family active site-containing proteins that catalyze the oxidation of thiols, reduction of disulfide bonds, and isomerization of disulfides, depending on the reaction conditions. PDI catalyzes the formation of correct disulfide pairing in nascent proteins. PDI preferentially interacts with peptides that contain cysteine residues but is otherwise undiscriminating. The broad substrate specificity of PDI enables it to 20 speed the folding of diverse disulfide-containing proteins. By shuffling disulfide bonds, PDI enables proteins to quickly find the most thermodynamically stable pairings amongst available cysteine residues. Consequently, PDI is involved in protein processing, protein folding, and protein secretion. PDI is also involved in collagen and collagen-like protein biosynthesis and mutations of PDI cause 25 Ehlers-Danlos syndrome.

[0751] An embodiment of the invention is directed to a composition comprising a PDI polypeptide sequence of SEQ ID NO:46.

[0752] A further embodiment of the invention is directed to a composition comprising a PDI polypeptide fragment having a biological activity of binding peptides with cysteine residues.

[0753] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence of SEQ ID NO:45 encoding a PDI polypeptide.

5 [0754] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence encoding a PDI polypeptide fragment having biological activity of binding peptides with cysteine residues.

[0755] A preferred aspect of the invention is a host cell recombinant for polynucleotides encoding a PDI polypeptide or a biologically active fragment thereof. An additional preferred aspect 10 is a host cell recombinant for polynucleotides capable of directing PDI expression. Preferably, the polynucleotides capable of directing PDI expression are located in the 5' regulatory region of the PDI gene. Further preferably, these polynucleotides are located within 500 base pairs of the PDI coding region. These polynucleotides preferably comprise a promoter sequence. Techniques known in the art for introducing polynucleotide sequences to endogenous sequences are described in U.S. Patent 15 5641670 and PCT WO9629411, which disclosures are hereby incorporated by reference in their entireties. PDI protein produced by said host cell may be used for in vitro detection and purification methods as well as diagnosis and in vivo applications.

[0756] A method of oxidizing thiols species comprising the step of: contacting a PDI polypeptide with peptides with reduced cysteine residues. Preferably, thiols are oxidized to disulfide 20 bonds. More preferably, the disulfide bonds are correctly isomerized. This method may be applied to recombinant production of protein sequences containing collagen/ collagen-like protein sequences.

[0757] A preferred aspect of the invention is a host cell recombinant for polynucleotides encoding a PDI polypeptide or a biologically active fragment thereof. An additional preferred aspect 25 is a host cell recombinant for polynucleotides capable of directing PDI expression. Preferably, the polynucleotides capable of directing PDI expression are located in the 5' regulatory region of the PDI gene. Further preferably, these polynucleotides are located within 500 base pairs of the PDI coding region. These polynucleotides preferably comprise a promoter sequence.

[0758] An embodiment of the present invention relates to compositions comprising PDI polypeptides. The method of producing PDI polypeptides comprises the steps of: i) transfecting a 30 mammalian host cell with a recombinant expression vector comprising a polynucleotide of the present invention, and ii) purifying the produced protein. The purification of the protein can be done following any technique well known to those skilled in the art. Preferably, an antibody directed

against PDI or part thereof, preferably, an antibody directed against the Thioredoxin domain of the polypeptide, may be bound to a chromatographic support to form an affinity chromatography column.

[0759] Another embodiment of the invention provides for a method utilizing PDI alone or in combination with a reductant or reduction system to reduce protein intramolecular disulfide bonds. 5 For example, PDI may be used in combination with glutenins or gliadins present in flour or seeds to improve dough strength and baked goods characteristics such as crumb quality, softness, and higher loaf volume as further described in US patent 6,113,951, which disclosure is hereby incorporated in its entirety.

[0760] Further preferred is a method of reducing a toxic protein having one or more 10 intramolecular cystines comprising contacting the toxic protein with an amount of PDI effective for reducing the protein, and maintaining the contact for a time sufficient to reduce one or more disulfide bridge and denature the toxic protein (as further described in US patents 6,113,951, which disclosure is hereby incorporated in its entirety).

[0761] In one embodiment, PDI may be co-expressed with a desired protein or proteins in a 15 host cell which is used to produce the desired protein or proteins. The co-expression of PDI will result in an increase in the amount of correctly folded protein obtained from the host cell. In this method, a first vector expressing PDI and a second vector expressing a desired protein are introduced into a host cell using conventional methods, and following expression, the desired protein is harvested.

[0762] In another embodiment, PDI may be added to a protein sample during processing to 20 increase the efficiency of protein processing. Processing of protein will be carried out following expression of at least one desired protein in the host cell and may include purification, renaturation, solubilization or any other steps used to obtain a pure and active form of a desired protein. In this embodiment, PDI may be added during these steps to minimize loss of protein resulting, for example, from protein oxidation, aggregation, or improper folding. The amount of PDI to be added may be 25 determined by one of skill using routine methods, and will be sufficient to reduce the loss of protein.

[0763] In another embodiment, the invention provides a method for neutralizing food allergens by PDI as described in US Patent 6190723, which disclosure is hereby incorporated by reference in its entirety. A method of decreasing the allergenicity of an allergenic food protein comprises the step of: contacting the protein with and amount of PDI, Nicotinamide adenine 30 dinucleotide phosphate-thioredoxin reductase and NADPH, or an amount of PDI and dithiothreitol effective for decreasing the allergenicity of the protein.

[0764] In another embodiment, the invention provides a method for increasing the digestibility of food proteins. This method may be carried out as described in US Patent 5,952,034, which disclosure is hereby incorporated by reference in its entirety. A method of increasing the digestibility of a food protein comprises the step of: treating a food with an amount of PDI, 5 nicotinamide adenine dinucleotide thioredoxin reductase (NTR) and NADPH effective for increasing the digestibility of the food.

[0765] An embodiment of the invention provides for a method of screening test substances for modulators of PDI expression. This method comprises the steps of: i) contacting a cell with a test substance; and ii) comparing PDI expression in the cell after exposure to the test substance to that of 10 an unexposed control cell. PDI expression is determined by methods common to the art or included herein, by detecting PDI polynucleotides or polypeptides. An example of this method comprises the steps of: i) culturing two equivalent cell samples; ii) adding a test substance to one of the cultures and not the other; iii) harvesting both cultures at a specified time; iv) purifying the mRNA from each sample of cells; v) comparing the level of PDI mRNA in each sample by Northern blot, RTPCR, or 15 another method common to the art. The invention provides for design and use of specific polynucleotide probes and primers, as discussed herein. An additional example comprises the steps of: i) having two equivalent cultures of cells; ii) adding a test substance to one of the cultures and not the other; iii) harvesting both cultures; iv) purifying the protein from each sample of cells; v) comparing 20 the level of PDI polypeptides in each sample by Western blot, immunohistochemistry, or another method common to the art. The invention provides for design and use of specific antibodies and antibody fragments, as discussed herein. This method may be applied to diagnosis of PDI-related disorders. For example, an abnormally high level of PDI indicates presence of primary tumors such as lung, colon, cervical and hepatocellular carcinoma.

[0766] Over-expression of PDI plays a role in the development of disease. Increased 25 expression of PDI is associated with Sjogrens' syndrome, increased growth of leukemia and lymphoma cells, and the proliferation of virally transformed cells. A deficiency of PDI is associated with immunological diseases, including, but not limited to, atherosclerosis, stroke, asthma, allergies, Crohns' disease, ulcerative colitis, diabetes, Hermansky-Pudlack syndrome, Alzheimers disease, and damage to tissues caused by trauma, ischemia, hypoxia, radiation and ultraviolet exposure.

30 [0767] In another embodiment, a PDI polypeptide or a fragment thereof, may be used to screen for compounds that activate or inhibit PDI activity. This method comprises the steps of: i) contacting a PDI polypeptide or fragment thereof with a test substance and ii) monitoring PDI activity.

- [0768] Test substances that decrease PDI expression or activity are defined as inhibitors or antagonists of PDI. Test substances that increase PDI expression or activity are defined as activators or agonists. Inhibitors of PDI include, but are not limited to antisense oligonucleotides, ribozymes, and antibodies. These agents may be made and used according to methods well known in the art.
- 5 Disorders characterized by aberrant PDI expression is a cardiovascular disorder, e.g. atherosclerosis, ischaemia reperfusion injury, cardiac hypertrophy, hypertension, coronary artery disease, myocardial infarction, arrhythmia, cardiomyopathies, and congestive heart failure; a connective tissue disorder, e.g., Ehlers-Danlos Syndrome; or a hepatic disorder, e.g., alcoholic liver disease, liver cirrhosis and liver cancer.
- 10 [0769] In another embodiment, the protein of the present invention, as well as agents, or modulators which have a stimulatory or inhibitory effect on PDI activity as identified by a screening assay can be administered to individuals to treat (prophylactically or therapeutically) disorders, described herein, associated with aberrant PDI activity.
- [0770] In another embodiment of the invention, PDI, its catalytic or immunogenic fragments 15 or oligopeptides thereof, can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes, between PDI and the agent being tested, may be measured. Another technique for drug screening which may be used provides for high throughput screening of compounds having suitable binding 20 affinity to the protein of interest as described in published PCT application WO84/03564, which disclosure is hereby included in its entirety. In this method, as applied to PDI, large numbers of different small test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. The test compounds are reacted with PDI, or fragments thereof, and washed. Bound PDI is then detected by methods well known in the art. Purified PDI can also be coated directly onto plates 25 for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

[0771] In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding PDI specifically compete with a test compound for binding PDI. In this manner, the antibodies can be used to detect the presence of any peptide which shares one 30 or more antigenic determinants with PDI.

[0772] In one embodiment, PDI or a biologically active fragment thereof is administered to a subject to treat conditions or diseases associated with a deficiency of PDI. These include, but are not

limited to, cancers of the brain, prostate, breast, bladder, and thyroid, and immunological and infectious diseases, including, but not limited to, atherosclerosis, stroke, asthma, allergies, Crohns' disease, ulcerative colitis, diabetes, Hermansky-Pudlack syndrome, Alzheimers disease, and damage to tissues caused by trauma, ischemia, hypoxia, radiation and ultraviolet exposure.

5 [0773] In another embodiment, a vector capable of expressing PDI, or a fragment or a derivative thereof, may also be administered to a subject to treat or prevent diseases or conditions described above.

[0774] In one embodiment, agonists of PDI may be administered to a subject to treat or prevent diseases or conditions described above.

10 [0775] In another embodiment, a vector expressing antisense of the polynucleotide encoding PDI may be administered to a subject to treat or prevent conditions or diseases associated with over-expression of PDI. Such conditions or diseases include lymphoma, leukemia, Sjogrens' syndrome, human T-lymphotropic virus, Epstein-Barr virus, human immunodeficiency virus. An example is described in published PCT application WO 99/38963, which disclosure is hereby included in its 15 entirety .

[0776] Another embodiment of the invention relates to composition and methods using polynucleotide sequences encoding the protein of the invention or part thereof to establish transgenic model animals (*D. melanogaster*, *M. musculus*), by any method familiar to those skilled in the art. By modulating *in vivo* the expression of the transgene with drugs or modifier genes (activator or 20 suppressor genes), animal models can be developed that mimic human hormone-dependent disorders such as cancers. These animal models would thus allow the identification of potential therapeutic agents for treatment of the disorders. In addition, recombinant cell lines derived from these transgenic animals may be used for similar approaches *ex vivo*.

[0777] In another embodiment, an array of oligonucleotides probes comprising the nucleotide sequence of PDI or fragments thereof can be constructed to conduct efficient screening of e.g., genetic mutations or deletion. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, and to develop and monitor the activities of therapeutic agents (see for example: 30 Chee, M. et al., Science, 274:610-614 (1996) which disclosure is hereby incorporated by reference in its entirety). For example, mutations of PDI cause Ehlers-Danlos syndrome.

Protein of SEQ ID NO:48 (internal designation Clone 146821_106-020-1-0-G3-F)

[0778] The cDNA of Clone 146821_106-020-1-0-G3-F (SEQ ID NO:47) encodes NBART protein of SEQ ID NO:48, comprising the amino acid sequence:

MLVMYLLAALFGYLTFYGEVEDELLHAYSKVYTLIDIPLLMVRALVAVTLPVILFPIRTS
VITLLFPKRPFSWIRHFLIAAVLIALNNVLVILVPTIKYIFGFIGASSATMLIFILPAVFYLKLVKK

5 ETFRSPQKVGVALIFLVVGIFFMIGSMALIIDWIYDPPNSKHH. Accordingly, it will be appreciated that all characteristics and uses of the polypeptides of SEQ ID NO:48 described throughout the present application also pertain to the polypeptides encoded by the nucleic acids included in Clone 146821_106-020-1-0-G3-F. In addition, it will be appreciated that all characteristics and uses of the polynucleotides of SEQ ID NO:47 described throughout the present application also pertain to the 10 nucleic acids comprising the human cDNA in Clone 146821_106-020-1-0-G3-F. A preferred embodiment of the invention is directed toward the compositions of SEQ ID NO:47, SEQ ID NO:48 and Clone 146821_106-020-1-0-G3-F. Also preferred are polypeptide fragments having a biological activity as described herein and the polynucleotides encoding the fragments.

[0779] ADP ribosylation factor (ARF) and ARF-like (ARL) proteins comprise the ARF 15 family within the Ras superfamily of regulatory GTPases. Members of this superfamily function as molecular nodes in signaling that can directly activate one or more enzymatic activities or coordinate the recruitment and assembly of more elaborate multi-subunit complexes. ARL2 is a member of a functionally distinct group of ARF-like genes. ARL2 binds GTP rapidly and hydrolyzes it. In the GDP-bound form, ARL2 interacts with the tubulin-specific chaperone cofactor D. This interaction 20 prevents the destruction of tubulin and microtubules by overexpressed cofactor D. The tubulin GTPase activating protein activity of cofactor D is also inhibited by ARL2 binding. This interaction prevents ARL2 contributes to modulating microtubule dynamics. NBART binds specifically to ARL2-GTP with high affinity but does not interact with ARL2-GDP.

[0780] An embodiment of the invention is directed to a composition comprising a NBART 25 polypeptide sequence of SEQ ID NO:48.

[0781] A further embodiment of the invention is directed to a composition comprising a NBART polypeptide fragment having a biological activity of binding to ARL2.

[0782] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence of SEQ ID NO:47 encoding a NBART polypeptide.

30 [0783] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence encoding a NBART polypeptide fragment having biological activity of binding to ARL2.

[0784] A preferred aspect of the invention is a host cell recombinant for polynucleotides encoding a NBART polypeptide or a biologically active fragment thereof. An additional preferred aspect is a host cell recombinant for polynucleotides capable of directing NBART expression. Preferably, the polynucleotides capable of directing NBART expression are located in the 5' regulatory region of the NBART gene. Further preferably, these polynucleotides are located within 500 base pairs of the NBART coding region. These polynucleotides preferably comprise a promoter sequence. Techniques known in the art for introducing polynucleotide sequences to endogenous sequences are described in U.S. Patent 5641670 and PCT WO9629411, which disclosures are hereby incorporated by reference in their entireties. NBART protein produced by said host cell may be used 10 for in vitro detection and purification methods as well as diagnosis and in vivo applications.

[0785] A preferred embodiment of the invention is a method of using NBART to bind ARL2. This method comprises the step of contacting a NBART polypeptide or active fragment thereof with an ARL2 protein under conditions that allow NBART binding, whereby binding inhibits the activity of ARL2.

[0786] A preferred embodiment of the invention is a method of purifying ARL2. This method comprises the steps of: contacting a NBART polypeptide or active fragment thereof with a ARL2 under conditions that allow binding; removing contaminants; and eluting the ARL2 with more stringent conditions. Preferably, the NBART peptide is immobilized on a solid or semi-solid matrix to facilitate washing of sample to remove contaminants.

[0787] A preferred embodiment of the inventions is a method of detecting ARL2. This method comprises the step of contacting a NBART polypeptide or active fragment thereof with ARL2 and detecting the presence of said ARL2 by detecting NBART. Preferably, the NBART polypeptide is detectably labeled with, for example, a fluorescent, luminescent, or radioactive compound. Preferably, ARL2 is detected in a biological fluid such as cell culture media and body fluids. This method may be applied to quantifying the level of ARL2 expression in a cell sample or individual. This information may be useful in determination of microtubule-related disorders, such as, but not limited to, cancer, 25 human polycystic kidney disease, Alzheimer disease, and Down syndrome.

[0788] An additional preferred aspect is a host cell recombinant for polynucleotides capable of directing NBART expression. Preferably, the polynucleotides capable of directing NBART expression are located in the 5' regulatory region of the NBART gene. Further preferably, these polynucleotides are located within 500 base pairs of the NBART coding region. These polynucleotides preferably comprise a promoter sequence.

[0789] A further aspect of the present invention involves the isolation and purification of NBART. An embodiment of the present invention relates to compositions comprising NBART polypeptides. The method of producing NBART polypeptides comprises the steps of: i) transfecting a mammalian host cell with a recombinant expression vector comprising a polynucleotide of the present invention, and ii) purifying the produced protein. NBART may be isolated according to any technique known in the art or disclosed herein. Preferably, an antibody directed against NBART or a fragment thereof, is bound to a chromatographic support to form an affinity chromatography column.

Protein of SEQ ID NO:50 (internal designation Clone 644724_181-21-1-0-A12-F)

[0790] The cDNA of Clone 644724_181-21-1-0-A12-F (SEQ ID NO:49) encodes NBTG 10 protein of SEQ ID NO:50, comprising the amino acid sequence:

MHPFYTRAATMIGEIAAAVSFISKFLRTKGLTSERQLQTFSQLQELLAEHYKHHWFPEKPC
GSGYRCIRINHKMDPLIGQAAQRIGLSSQELFRLLPSELTLWVDPYEVSYRIGEDGSICVLYEAS
PAGGSTQNSTNVQMVDSDRISCKEELLGRTPSKNYNMMTVSS. Accordingly, it will be
appreciated that all characteristics and uses of the polypeptides of SEQ ID NO:50 described
15 throughout the present application also pertain to the polypeptides encoded by the nucleic acids
included in Clone 644724_181-21-1-0-A12-F. In addition, it will be appreciated that all characteristics
and uses of the polynucleotides of SEQ ID NO:49 described throughout the present application also
pertain to the nucleic acids of the human cDNA in Clone 644724_181-21-1-0-A12-F. A preferred
embodiment of the invention is directed toward the compositions of SEQ ID NO:49, SEQ ID NO:50
20 and Clone 644724_181-21-1-0-A12-F. Also preferred are polypeptide fragments having a biological
activity as described herein and the polynucleotides encoding the fragments.

[0791] NBTG is a novel polymorphism variant of B-cell translocation gene 1 protein (accession numbers P31607). The protein of the invention displays an anti-proliferative domain:

IGEIAAAVSFISKFLRTKGLTSERQLQTFSQLQELLAEHYKHHWFPEKPCKGSGYRCIRINHK
25 MDPLIGQAAQRIGLSSQELFRLLPSELTLWVDPYEVSYRIGEDGSICVLYEASPAGGSTQNSTN
VQMVDSDRISCKEELLGRTPSKNYNMMTVSS. Accordingly, some embodiments of the present
invention relate to polypeptides comprising the anti-proliferative domain.

[0792] The ability to negatively regulate cell proliferation is a necessity for all living organisms. Unicellular organisms must limit their replication to the time when adequate nutrients and
30 other environmental factors are present, and multicellular organisms must accurately shape and maintain the architecture of their component tissues. The failure in a multicellular organism to provide adequate negative growth control in the developmental period may result in a malformation, which

may be lethal. In the postdevelopmental period, such a failure may result in neoplasia. Because negative control is so critical, specific genes have evolved whose role is actively antiproliferative.

[0793] NBTG is a member of a family of antiproliferative genes. NBTG is crucial to counteract the growth inducing elements and have the same importance as proto-oncogenes in controlling cell division. Loss of NBTG is associated with irregular cellular differentiation and proliferation or with alteration of embryonic development. NBTG downregulates N1H3T3 cell proliferation when over expressed. NBTG gene is involved in a chromosomal translocation in B-cell chronic lymphocytic leukemia. NBTG is expressed in tissues (lymphoid, liver, plasma) containing non-dividing cells likely to re-enter the cell cycle upon stimuli, however, NBTG is barely detectable in fully differentiated tissues such as brain and muscle.

[0794] An embodiment of the invention is directed to a composition comprising a NBTG polypeptide sequence of SEQ ID NO:50.

[0795] A further embodiment of the invention is directed to a composition comprising a NBTG polypeptide fragment having a biological activity of inhibiting cell proliferation.

[0796] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence of SEQ ID NO:49 encoding a NBTG polypeptide.

[0797] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence encoding a NBTG polypeptide fragment having a biological activity of inhibiting cell proliferation.

[0798] A further embodiment of the invention is directed to a composition comprising an antibody directed against the NBTG polypeptide or a NBTG polypeptide fragment having biological activity. Preferably, the antibody recognizes a non-linear epitopes, and specifically binds to the C-terminal sequence of NBTG.

[0799] A preferred aspect of the invention is a host cell recombinant for polynucleotides encoding a NBTG polypeptide or a biologically active fragment thereof. An additional preferred aspect is a host cell recombinant for polynucleotides capable of directing NBTG expression. Preferably, the polynucleotides capable of directing NBTG expression are located in the 5' regulatory region of the NBTG gene. Further preferably, these polynucleotides are located within 500 base pairs of the NBTG coding region. These polynucleotides preferably comprise a promoter sequence. Techniques known in the art for introducing polynucleotide sequences to endogenous sequences are described in U.S. Patent 5641670 and PCT WO9629411, which disclosures are hereby incorporated by

reference in their entireties. NBTG protein produced by said host cell may be used for in vitro detection and purification methods as well as diagnosis and in vivo applications.

[0800] An embodiment of the present invention relates to compositions comprising NBTG polypeptides. The method of producing NBTG polypeptides comprises the steps of: i) transfecting a 5 mammalian host cell with a recombinant expression vector comprising a polynucleotide of the present invention, and ii) purifying the produced protein. The purification of the protein can be done following any technique well known to those skilled in the art. Preferably, an antibody directed against NBTG or part thereof, more preferably, an antibody directed against the C-terminal sequence of NBTG polypeptide, may be bound to a chromatographic support to form an affinity 10 chromatography column.

[0801] In another embodiment, the invention provides methods and compositions for detecting the level of expression of NBTG mRNA. Quantification of mRNA levels of NBTG polypeptides is useful for the diagnosis or prognosis of diseases associated with an altered expression of the protein of the invention. Conditions, diseases or disorders associated with altered expression 15 include, but are not limited to, aberrant cellular proliferation, such as cancer, psoriasis, blood vessel proliferative disorders, fibrotic disorders, and actinic lesions. Assays for the detection and quantification of the mRNA of the protein of the invention are well known in the art (see, for example, Maniatis, Fitsch and Sambrook, Molecular Cloning; A Laboratory Manual (1982), or Current Protocols in Molecular Biology, Ausubel, F.M. *et al.* (Eds), Wiley & Sons, Inc.). For example, the 20 nucleic acid molecule or probe may be labeled by standard methods and added to a biological sample from a patient under conditions of formation of hybridization complexes. After an incubation period, the sample is washed and the amount of label (or signal) associated with hybridization is quantified and compared with a standard value. If the amount of label in the patient sample is significantly 25 altered in comparison to the standard value, then the presence of the associated condition, disease or disorder is indicated. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies and in clinical trial and to monitor the treatment of an individual patient. Once the presence of a condition is established and a treatment protocol is initiated, diagnostic assays may be repeated on a regular basis to determine if the level of expression in the patient begins 30 to approximate that which is observed in a normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

[0802] In another embodiment, the invention relates to methods and compositions for detecting and quantifying the level of NBTG present in a particular biological sample. These methods

are useful for the diagnosis or prognosis of diseases associated with altered levels of the protein of the invention. Diagnostic assays to detect NBTG may require cells from a biopsy, in situ assay of cells from organ or tissue sections, or an aspirate of cells from a tumor or normal tissue. In addition, assays may be conducted upon cellular extracts from organs, tissues, cells, urine, or serum or blood or any other body fluid or extract. Detection and quantification of NBTG polypeptides using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunoabsorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS).

[0803] NBTG has an antiproliferative ability, and thus may be employed to treat diseases or pathological conditions associated with aberrant cellular proliferation and malignant conditions. The polypeptides may be employed to inhibit tumor growth and cell proliferation. NBTG may also be employed to prevent scar formation at the site of wound healing. Restenosis, or re-occlusion of arterial walls after balloon angioplasty, may also be treated with NBTG as arteries re occlude through cell proliferation. Similarly angiogenesis into a tumor may be inhibited.

[0804] In one embodiment, NBTG or a biologically active fragment thereof is administered to treat or prevent a condition associated with altered NBTG expression or activity. Examples of such conditions include, but are not limited to, those described above. Most preferably, a composition comprising a NBTG polypeptide is administered to an individual suffering from abnormal or undesirable cell proliferation, such as, e.g., tumor growth, endothelial cell proliferation, and angiogenesis related to tumor growth. Tumors that can be treated with the compositions of the present invention include, but are not limited to, gliomas, amelanotic melanomas, prostate tumors and lung tumors.

[0805] In another embodiment, a pharmaceutical composition comprising a substantially purified NBTG polypeptide in conjunction with a pharmaceutical carrier may be administered to a subject to treat or prevent a condition associated with altered expression or activity of the endogenous protein including, but not limited to, those provided above.

[0806] In another embodiment, a substance which modulates the activity of NBTG may be administered to a subject to treat or prevent a condition associated with altered lifespan, expression, or activity of NBTG polypeptides including, but not limited to, those listed above. In one aspect, an antibody which specifically binds NBTG may be used as a targeting and delivery mechanism for bringing a pharmaceutical agent to cells or tissue which express NBTG like, but not limited to, the liver cells.

[0807] Another embodiment of the subject invention provides compositions and methods of selectively increasing the activity of the protein of the invention. Activation of NBTG allows for the successful treatment and/or management of diseases or biochemical abnormalities associated with NBTG activity. Agonists, able to increase the expression or the activity of the protein of the invention, 5 are useful in the treatment of diseases associated with cell proliferation, like, but not limited to, cancer, psoriasis, blood vessel proliferative disorders, fibrotic disorders, and actinic lesions. Alternatively, antagonist, able to decrease the expression or the activity of NBTG, are useful in the treatment of conditions characterized by insufficient cellular proliferation. The conditions to be treated include, for example, osteoporosis, fragile skin and poor wound healing.

[0808] In an additional embodiment, a vector capable of expressing NBTG or a preferred fragment may be administered to a subject to treat or prevent a condition associated with altered lifespan, expression, or activity of the protein of the invention including but not limited to, those listed above. NBTG gene and gene products may also be employed for modulation of cellular growth. Due to their anti-proliferative effect they could be selectively administered or possibly inhibited when it is 15 desirable to have certain cells proliferate. An example would be a disorder related to the underproduction of certain cells, where proliferation and differentiation of these cells would help to treat the disorders listed above.

[0809] In some embodiments, the invention also concerns a diagnostic kit for detecting *in vitro* the presence of NBTG polypeptide. This kit comprises: a polyclonal or monoclonal antibody or 20 fragment thereof that specifically binds a NBTG polypeptide; and optionally, ii) a reagent allowing the detection of the antigen-antibody complexes formed. Preferably, the antibody, or antibody fragment is detectably labeled. Such labels include fluorescent luminescent, and radioactive compounds, as well as enzymatic substrates. The optional reagent may provide a detectable signal in either bind to the antibody or react with the label on the antibody. NBTG antibodies may be used to diagnose low 25 proliferative diseases such as osteoporosis, fragile skin and poor wound healing. To diagnose such disorders, an appropriate biological sample can be tested to determine the level of NBTG being produced.

[0810] In another embodiment, the current invention provides a method of effectively blocking proliferation or inhibiting the growth of a cell *in vivo* or *in vitro*. Preferably, the present 30 invention can be used to stop unrestrained cell proliferation and to eliminate as many tumor cells as possible. The method will be performed by administering an antisense oligonucleotide directed against NBTG to a cell. One strategy for delivering antisense oligonucleotides to targeted cells involves encapsulation or incorporation of the therapeutic bioactive molecules in liposomes, such as

cationic liposomes. These liposomes are known to provide a shield against nucleotide degradation *in vivo* and can be targeted to specific areas of the body at which point they slowly release their contents. Alternatively, a polynucleotide construct comprising plasmid DNA operably linked to the antisense oligonucleotide could be used. The nucleotide sequences are administered *in vivo* in a suitable buffer 5 or carrier solution known to those skilled in the art.

[0811] Another embodiment of the invention relates to composition and methods using polynucleotide sequences encoding the protein of the invention or a fragment thereof to establish transgenic model animals (*D. melanogaster*, *M. musculus*), by any method familiar to those skilled in the art. By modulating *in vivo* the expression of the transgene with drugs or modifier genes (activator 10 or suppressor genes), animal models can be developed that mimic human hormone-dependent disorders such as cancers. These animal models would thus allow the identification of potential therapeutic agents for treatment of the disorders. In addition, recombinant cell lines derived from these transgenic animals may be used for similar approaches *ex vivo*.

Protein of SEQ ID NO:52 (Internal designation 583702_181-8-4-0-C8-F)

15 [0812] The cDNA of 583702_181-8-4-0-C8-F (SEQ ID NO:51) encodes protein vITH1 of SEQ ID NO:52, comprising the amino acid sequence:

MPLPLPSAFVLSALQPSPTHSSSNTQRLPDRVTGGFSVNGQLIGNKARSPGQHDGTYFGR
LGIANPATDFQLEVTPQNITLNPGFGGPVFSWRDQAVLRQDGVVVTINKKRNLVSVDDG
GTFEVVLHRVWKGSVHQDFLGFYVLDSHRMSARTHGLGQFFHPIGFEVSDIHPGSDPT

20 KPDATMVVRNRRLTVTRGLQKDYSKDPWHGAEVSCWFIHNNNGAGLIDGAYTDYIVPDIF.

Accordingly, it will be appreciated that all characteristics and uses of polypeptides of SEQ ID NO:52 described throughout the present application also pertain to the polypeptides encoded by the nucleic acids comprising the human cDNA in Clone 583702_181-8-4-0-C8-F. In addition, it will be appreciated that all characteristics and uses of the polynucleotides of SEQ ID NO:51 described 25 throughout the present application also pertain to the nucleic acids included in 583702_181-8-4-0-C8-F. A preferred embodiment of the invention is directed toward the compositions of SEQ ID NO:51, SEQ ID NO:52, and 583702_181-8-4-0-C8-F. Also preferred are polypeptide fragments having a biological activity as described herein and the polynucleotides encoding the fragments.

[0813] The cDNA of SEQ ID NO:51 is a splice variant of the heavy chain 1 of the human 30 inter-alpha-trypsin inhibitor (ITH1), encoding a 239 amino-acid protein of SEQ ID NO:52 named vITH1.

[0814] Inter-alpha-trypsin inhibitor belongs to the superfamily of Kunitz-type serine protease inhibitors in mammals. It is a glycosylated protease inhibitor with a polypeptide chain structure composed of two heavy chains (HC1 and HC2) covalently linked to a light chain (bikunin). Bikunin functions as the protease inhibitor. vITH1 polypeptide is a novel splice variant of the heavy chain 1 that lacks the bikunin linkage site. However, ITH1 binds hyaluronic acid (HA) protein. The connective tissues are the main mammalian source of HA but it is also found in the extracellular matrix, cartilage, bone marrow, synovial fluid. After endocytosis, HA is normally catabolized by hyaluronidase digestion in the liver.

[0815] An embodiment of the invention is directed to a composition comprising a vITH1 polypeptide sequence of SEQ ID NO:52.

[0816] A further embodiment of the invention is directed to a composition comprising a vITH1 polypeptide fragment having a biological HA-binding protein activity.

[0817] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence of SEQ ID NO:51 encoding a vITH1 polypeptide.

[0818] A further embodiment of the invention is directed to a composition comprising a polynucleotide sequence encoding a vITH1 polypeptide fragment having a biological HA-binding protein activity.

[0819] A further embodiment of the invention is directed to a composition comprising an antibody directed against a vITH1 polypeptide sequence of SEQ ID NO: 52 or a vITH1 polypeptide fragment having a biological HA-binding protein activity. Preferably, the antibody specifically binds to vITH1 but not to ITH1. Preferably, the antibody specifically recognizes an epitope comprising the amino acids: VTGG, TGGF, GGFS, VTGGFS, MPLP, PLPL, PLPS, LPSA, or MPLPLPSA.

[0820] An embodiment of the present invention relates to compositions comprising vITH1 polypeptides. The method of producing vITH1 polypeptides comprises the steps of: i) transfecting a mammalian host cell with a recombinant expression vector comprising a polynucleotide of the present invention, and ii) purifying the produced protein. The purification of the protein can be done following any technique well known to those skilled in the art. Preferably, an antibody directed against vITH1 or fragment thereof, as described above, may be bound to a chromatographic support to form an affinity chromatography column.

[0821] In another embodiment, the invention is directed to a method of detecting vITH1 polypeptide in a biological sample, said method comprising the steps of: i) contacting a biological

sample with an antibody or antibody fragment that specifically binds vITH1 polypeptide; and ii) detecting the antigen-antibody complex formed. The antibody or antibody fragment may be monoclonal or polyclonal. In addition, the antibody or antibody fragment may be primarily or secondarily labeled by a detectable compound (e.g., radioactive, fluorescent, luminescent, or 5 enzymatic) common in the art.

[0822] In a further embodiment, vITH1 polypeptide is used to purify HA. In such method, vITH1 polypeptide is preferably covalently or non-covalently attached to a solid matrix and allowed to bind HA using techniques well known in the art. This method comprises the steps of: i) washing the solid matrix to get rid of contaminants, ii) eluting the particle of interest using more stringent 10 conditions. Additional aspects of this embodiment include methods of using vITH1 polypeptide to detect and quantify HA using techniques common in the art. This method comprises the steps of: i) obtaining a biological sample suspected of containing HA, ii) contacting said sample with a vITH1 polypeptide or fragment thereof under conditions suitable for binding of vITH1 and detecting the presence or absence of HA by detecting vITH1. Preferably, vITH1 polypeptide or fragment thereof is 15 covalently attached to a detectable compound. Alternatively, a detectable vITH1-specific antibody or fragment thereof may be used to detect vITH1. This embodiment is useful, for example, as a diagnostic tool for detecting the presence of HA for individuals at risk of or suffering from a wide range of diseases associated with an abnormal HA storage. For example, it can be applied in the diagnosis of conditions such as, but not limited to, hyaluronidase deficiency also known as type IX 20 mucopolysaccharidosis, disorders of HA metabolism associated with generalized folding and thickening of the skin as described by Ramsden CA et al, J Pediatr 2000 Jan;136(1):62-68), liver diseases such as cirrhosis, renal failure and several lung diseases, arthritis diseases such as rheumatoid arthritis and osteoarthritis, and certain cancers such as mesothelioma and Wilms' tumors. Such method is also useful to evaluate the severity degree of such diseases. This kit comprises: a polyclonal 25 or monoclonal antibody or fragment thereof that specifically binds a vITH1 polypeptide; and optionally, ii) a reagent allowing the detection of the antigen-antibody complexes formed. Preferably, the antibody or antibody fragment is detectably labeled. Such labels include fluorescent, luminescent, and radioactive compounds, as well as enzymatic substrates. The optional reagent may provide a detectable signal and either bind to the antibody or react with the label on such antibody. To 30 determine the level of HA and diagnose or manage such disorders, preferred appropriate biological samples are serum, urine and synovial fluid in the case of arthritis disorders.

[0823] An embodiment of the invention provides for a method of screening test substances for modulators of vITH1 expression. This method comprises the steps of: i) contacting a cell with a test

substance, and ii) comparing vITH1 expression in the cell after exposure to the test substance to that of an unexposed control cell. vITH1 expression is determined by methods common to the art or included herein, by detecting vITH1 polynucleotides or polypeptides. An example of this method comprises the steps of: i) culturing two equivalent cell samples, ii) adding a test substance to one of the cultures and 5 not the other, iii) harvesting both cultures at a specified time, iv) purifying the mRNA from each sample of cells, v) comparing the level of vITH1 mRNA in each sample by Northern blot, RTPCR, or another method common to the art. The invention provides for design and use of specific polynucleotide probes and primers, as discussed herein. An additional example comprises the steps of: i) having two equivalent cultures of cells, ii) adding a test substance to one of the cultures and not the 10 other, iii) harvesting both cultures, iv) purifying the protein from each sample of cells, v) comparing the level of vITH1 polypeptides in each sample by Western blot, immunohistochemistry, or another method common to the art. The invention provides for design and use of specific antibodies and antibody fragments, as discussed herein.

[0824] In another embodiment, a vITH1 polypeptide or a fragment thereof, may be used to 15 screen for compounds that activate or inhibit vITH1 activity. This method comprises the steps of: i) contacting a vITH1 polypeptide or fragment thereof with a test substance and ii) monitoring vITH1 activity. vITH1 binds to HA. Thus, vITH1 activity may be monitored upon addition of the test substance by competitive binding assays with HA. In this aspect of the invention, a vITH1 polypeptide or fragment thereof may be free in solution, affixed to a solid support, recombinantly 20 expressed on or chemically attached to a cell surface, or located intracellularly. The formation of binding complexes between vITH1 polypeptide and the compound being tested, may be measured by methods well known to those skilled in the art, such as the BIACore (Upsala, Sweden). Another technique provides for high throughput screening of compounds having suitable binding affinity to the protein of the invention. Test substances that decrease vITH1 expression or activity are defined as 25 inhibitors or antagonists of vITH1. Test substances that increase vITH1 expression or activity are defined as activators or agonists. Agents which modulate the expression or activity of the vITH1 of the subject invention include, but are not limited to antisense oligonucleotides, ribozymes, and antibodies. These agents may be made and used according to methods well known in the art.

[0825] In another embodiment of the invention, the vITH1 polypeptide, fragments thereof, or 30 vITH1 agonists are used to bind HA *in vivo* and remove this molecule from the bloodstream. This method can be used to prevent or treat disorders associated with abnormal storage of HA, including those due to a decrease hepatic clearance, an excessive synthesis or an increased hyaluronidase activity, For example, vITH1 polypeptides, fragments thereof, or vITH1 agonists can be used in a

composition delivered to an individual at risk of or suffering from, but not limited to, hyaluronidase deficiency also known as type IX mucopolysaccharidosis, disorders of HA metabolism associated with generalized folding and thickening of the skin as described by Ramsden CA et al, J Pediatr 2000 Jan;136(1):62-68), liver diseases such as cirrhosis, renal failure and several lung diseases. In this embodiment, such compositions can also be used in any other conditions associated with elevated serum levels of HA such as arthritis diseases such as rheumatoid arthritis and osteoarthritis, and certain cancers such as, but not limited to, mesothelioma and Wilms' tumors. This method comprises the step of: introducing an effective amount of vITH1 polypeptide, fragment thereof or vITH1 agonist to the bloodstream of an individual. In this embodiment, the vITH1 polypeptide may further be expressed as a fusion protein with a polypeptide signal specifying excretion from the body. A preferred method of delivering vITH1 polypeptides or biologically active fragments thereof to an individual includes direct, intravenous injection of said polypeptides or fragments in a physiologically acceptable solution (e.g., pH-buffered isotonic saline solutions, pH-buffered isotonic saline solutions modified by addition of viscous elements such as glycerol). Alternatively, vITH1 polynucleotides may be introduced to express vITH1 polypeptides in the bloodstream. This method comprises the steps of: i) constructing a recombinant viral vector corresponding to a portion of the genome of an adenovirus capable of infecting a cell operatively linked to the nucleotide sequence of the invention and a regulatory sequence directing its expression; ii) delivery of an effective amount of the recombinant adenoviral vector to an individual with or at risk of suffering diseases or conditions mentioned above.

[0826] Another embodiment of the invention relates to animal models generated by modulating the expression or activity of the present protein in one or more tissues of the animal. These animals can be generated with any method of targeting overexpression or inactivation of vITH1, or by modulating the *in vivo* expression of the transgene with drugs or modifier genes (activator or suppressor genes). Such animals are useful for a number of purposes, because they represent an *in vivo* assay method for testing therapeutic candidate molecules potentially useful for the treatment of various pathophysiological aspects of diseases specifically related to the activity of vITH1. In addition, recombinant cell lines derived from these transgenic animals may be used for similar approaches *ex vivo*.

Nested N-Terminal Fragments of Related Proteins Acrp30R1L and Acrp30R1

[0827] The cDNA of Accession No. AAB30232 comprising the nucleotide sequence:

GGAAAACATGCCTGGGGCCGACGCTCTGCCCGGCTGCTGCCGCTGAGGAAAGCCGGGA
CGCGGAGCCCCGCCGAGAGCTTCTTGCTCCGGACGCCCTGGACGTGGCGGGCAGCCGC

GAGGGTAACCACCATGATCCCCTGGGTGCTCCTGGCCTGTGCCCTCCCTGTGCTGAC
CCACTGCTTGGCGCCTTGCTCGCAGGGACTCCGGAAAGGCTCCCTCAACTGGTCTGCA
GCCTGCCTGGCCCCCAGGGCCCACCCGGCCCCCAGGAGCCCCAGGGCCCTCAGGAATGA
TGGGACGAATGGGCTTCCTGGCAAAGACGGCCAAGATGGACACGACGGCACCAGGGGG
5 GACAGCGGAGAGGAAGGTCCACCTGGCCGGACAGGTACCCGGAAAGCCAGGACCAA
AGGGCAAAGCCGGGCCATTGGCGGGCTGGCCCCGTGGCCCAAGGGGTCAACGGT
ACCCCCGGGAAGCATGGCACACCAGGAAGAAGGGCCAAGGGCAAGAAAGGGGAGC
CAGGCCTCCCAGGCCCCCTGCAGCTGTGGCAGTGGCCATACCAAGTCAGCTTCTCGGTGG
CAGTGACCAAGAGCTACCCACGGGAGCAGCTGCCATCAAGTTGACAAGATTCTGATGA
10 ACGAGGGTGGCCACTACAATGCTCCAGCGGAAGTCGTCTGGCGTGCTGGATCT
ACTACTTCACCTACGACATCACGCTGGCCAACAAGCACCTGGCATCGGCCTGGTGCACA
ACGGCCAGTACCGCATCCGGACCTTGATGCCAACACCGGAACCAAGCAGATGTGGCCTCAG
GCTCCACCACCTGGCTCTCAAGCAGGGTACGAAGTTGGCTGCAGATCTTACTCAG
AGCAGAACAGGGCTTCTATGACCTTACTGGACAGACAGCCTTTACGGCTTCAAT
15 CTATGCCGACCAGGATGACCCAACGAGGTATAGACATGCCACGGCGGTCCAGGCA
GGGAACAAGCTCTGGACTTGGCTTACAGAGCAAGACCCACAACACTGTAGGCTGGGG
TGGGGGGTCGAGTGAGCGGTTCTAGCCTCAGGCTCACCTCCGCCTTTTTTCCCT
TCATTAAATCCAAACCTTTTATTCA

[0828] Encodes the protein Acrp30R1L comprising the amino acid sequence:

20 MIPWVLLACALPCAADPLLGFARRDFRKGPQLVCSLPGPQGPPGPPGAPGPSGMMGRMGF
PGKDGQDGHDGRGDSGEEGPPGRTGNRGKPGPKGKAGAIGRAGPRGPKGVNTPGKHGTP
GKKGPKGKKGEPGLPGPCSCSGHTKSAFSVAVTKSYPRERLPIKFDKILMNEGGHYNASSGK
FVCVGPGIYYFTYDITLANKHLAIGLVHNGQYRIRTFDANTGNHDVASGSTILALKQGDEVWL
QIFYSEQNGLFYDPYWTDSLFTGFLIYADQDDPNEV.

25 [0829] Accordingly, it will be appreciated that all characteristics and uses of polypeptides of the aforesaid Acrp30R1L amino acid sequence described throughout the present application also pertain to the polypeptides encoded by the nucleic acids included in the aforesaid Acrp30R1L nucleotide sequence. Also preferred are fragments having a biological activity as described therein and the polynucleotides encoding the fragments.

30 [0830] Acrp30R1L is comprised of at least three regions including:

1. an N-terminal putative signal sequence about from amino acids 1-20;
2. a unique region about from amino acids 21-39;

3. a collagen-like region about from amino acids 40-141; and
4. a globular region of C1q homology about from amino acids 149-285.

[0831] Acrp30R1L is comprised of several notable putative sites for proteolytic cleavage including:

- 5 1. collagenase matrix metalloproteinase-1 (MMP-1) between amino acids 46-47;
2. plasmin between amino acids 125-126, 126-127, 131-132, and 132-133; and
3. precerebellin processing protease between amino acids 133-134.

[0832] The cDNA of Accession No. AAZ45688 comprising the nucleotide sequence

GAATTCCGCACGAGGGCCCGAGGGTAACCACCATGATCCCCTGGGTGCTCCTGGCCTGT
10 GCCCTCCCTGTGCTGCTGACCCACTGCTTGGCGCCTTGCTCGCAGGGACTTCCGGAAAG
GCTCCCTCAACTGGTCTGCAGCCTGCCTGGCCCCAGGGCCCACCCGGCCCCCAGGAG
CCCCAGGGCCCTCAGGAATGATGGGACGAATGGGCTTCCTGGCAAAGACGGCCAAGAT
GGACACGACGGCGACCGGGGGGACAGCGGAGAGGAAGGTCCACCTGGCCGGACAGTGA
CCAAGAGCTACCCACGGGAGCGGCTGCCATCAAGTTGACAAGATTCTGATGAACGAG
15 GGTGGCCACTACAATGCTTCCAGCGGCAAGTCGTCAGCGGTGCCTGGATCTACTAC
TTCACCTACGACATCACGCTGGCCAACAAGCACCTGGCCATGGCCTGGTGCACAACGGC
CAGTACCGCATCCGGACCTTGATGCCAACACCGGCAACCACGATGTGGCCTCAGGCTCC
ACCATCCTGGCTCTCAAGCAGGGTACGAAGTTGGCTGCAGATCTCTACTCAGAGCAG
AACGGGCTTCTATGACCTTACTGGACAGACAGCCTTTACGGGCTTCAATCTATG
20 CCGACCAGGATGACCCAACGAGGTATAGACATGCCACGGCGGTCCAGGCAGGGAA
CAAGCTTCTGGACTTGGCTTACAGAGCAAGACCCCACAACACTGTAGGCTGGGGTGGGG
GGTCGAGTGAGCGGTTAGCCTCACGCTCACCTCTGCCTCTTTCCCCTTCATTA
AATCCAAACCTTTATTCAAAAAAAAAAAAAAGATGCGGCCG

[0833] Encodes the protein Acrp30R1 comprising the amino acid sequence

25 MIPWVLLACALPCAADPLLGFARRDFRKGSPQLVCSLPGPQGPPGPPGAPGPSGMGRMGF
PGKDGQDGHDGRGDSGEEGPPGRTVKSYPRERLPIKFDKILMNEGHHYNASSGKFVCGVVP
GIYYFTYDITLANKHLAIGLVHNGQYRIRTFDANTGNHDVASGSTILALKQGDEVWLQIFYSE
QNGLFYDPYWTDSLFTGFLIYADQDDPNEV.

[0834] Accordingly, it will be appreciated that all characteristics and uses of polypeptides of
30 the aforesaid Acrp30R1 amino acid sequence described throughout the present application also pertain

to the polypeptides encoded by the nucleic acids included in the aforesaid Acrp30R1 nucleotide sequence. Also preferred are fragments having a biological activity as described therein and the polynucleotides encoding the fragments.

[0835] Acrp30R1 is comprised of at least three regions including:

- 5 1. an N-terminal putative signal sequence about from amino acids 1-20;
 2. a unique region about from amino acids 21-39;
 3. a collagen-like region about from amino acids 40-87; and
 4. a globular region of C1q homology about from amino acids 88-217.

[0836] Acrp30R1 retains the putative site for MMP-1 cleavage present in Acrp30R1L but 10 lacks the putative sites in Acrp30R1L for cleavage by plasmin and precerebellin processing protease.

[0837] Acrp30R1L and Acrp30R1 are related proteins resulting from differential splicing of a single gene. Acrp30R1L and Acrp30R1 are comprised of a collagen domain and a C1q homology domain. Acrp30R1 lacks an internal 68 amino acid segment present in Acrp30R1L. The N-terminal fragments of Acrp30R1L and Acrp30R1 generated *in vivo* by proteolytic cleavage of the full-length 15 protein have unexpected and novel function as described below.

[0838] Chemokines are chemotactic cytokines that signal through G protein-coupled receptors (GPCR). Secondary lymphoid-tissue chemokine (SLC) is a polypeptide of 134 amino acids including a putative signal sequence of 23 amino acids (Nagira, M et al., J Biol Chem 272:19518-24 (1997) which disclosure is hereby incorporated by reference in its entirety). SLC is highly expressed by high 20 endothelial venules, specialized vessels involved in the homing of lymphocytes from the blood into lymph nodes and Peyer's patches (Murphy, PM et al., Pharmacological Reviews 52:145-176 (2000); Dieu-Nosjean, MC et al., J Leukoc Biol 66:252-62 (1999) which disclosures are hereby incorporated by reference in their entirety). SLC is chemotactic for T cells, B cells, and mature dendritic cells. SLC functions as an important mediator of physiological lymphocyte recirculation *in vivo* (Gunn, MD 25 et al, Proc Natl Acad Sci USA 95:258-63 (1998); Gunn, MD et al., J Exp Med 189:451-60 (1999) which disclosures are hereby incorporated by reference in their entirety). Recently it was shown that neurodegenerative stress induces neuronal expression of SLC and that SLC acts on microglia (Biber, K et al., Glia 34:121-33 (2001) which disclosure is hereby incorporated by reference in its entirety).

[0839] SLC binds to chemokine receptor CCR7 (Yoshida, R et al., J Biol Chem 273:7118-22 30 (1998) which disclosure is hereby incorporated by reference in its entirety). EBI1-ligand chemokine (ELC) also binds to CCR7. Recently it was shown that SLC also binds to the recently identified

chemokine receptor CCR10 (Gosling, J et al., J Immunol 164:2851-6 (2000) which disclosure is hereby incorporated by reference in its entirety). SLC and ELC bind competitively and only to CCR7 and CCR10.

[0840] SLC has been implicated in chronic inflammation, including rheumatoid arthritis
5 (Patel, DD et al., Clin Immunol 99:43-52 (2001) which disclosure is hereby incorporated by reference in its entirety) and prediabetic NOD mice (Hjelmstrom, P et al., Am J Pathol 156:1133-8 (2000) which disclosure is hereby incorporated by reference in its entirety). SLC (and ELC) enhance the replication
of HIV-1 in secondary lymphoid tissues (Nagira, M et al., Virology 264:422-6 (1999) which disclosure
is hereby incorporated by reference in its entirety). SLC has been shown to suppress proliferation of
10 myeloid progenitor cells (Kim, CH et al., J Leukoc Biol 66:455-61 (1999) which disclosure is hereby
incorporated by reference in its entirety). ELC (but not SLC) has been implicated in human
atheriosclerosis (Reape, TJ et al., Am J Pathol 154:365-74 (1999) which disclosure is hereby
incorporated by reference in its entirety).

[0841] SLC and ELC have been found to facilitate anti-tumor responses *in vivo*. SLC was
15 found to enhance the anti-tumor response in several mouse models (Kirk, CJ et al., Cancer Res
61:2062-70 (2001); Nomura, T et al., Int J Cancer 91:597-606 (2001); Sharma, S et al., J Immunol
164:4558-63 (2000) which disclosures are hereby incorporated by reference in their entirety). ELC
was found to mediate tumor rejection of murine breast cancer cells (Braun, SE et al., J Immunol
164:4025-31 (2000) which disclosure is hereby incorporated by reference in its entirety).

[0842] Acrp30R1L and Acrp30R1 are characterized by an amino acid motif that is otherwise
20 specific for chemokine SLC, as determined by BLAST analysis of the public protein database. The
specificity of this motif for SLC extends across species. The motif, encompassing amino acids 21 to
46 of Acrp30R1L and Acrp30R1 (and therefore N-terminal to the putative protease cleavage sites
within Acrp30R1L and Acrp30R1) is: RxxRKxxPxLxCsXP (where “x” is an unassigned amino acid).
25 This specific motif is conserved between human (Accession Nos. AAY12316, O00585, AAG03773,
AAW87589), mouse (O09006, AAG45834) and pig (AAW50886) SLC (encompassing amino acids 46
to 60 of SLC). This motif is inferred to be important for SLC function, namely for binding to
chemokine receptors CCR7 and CCR10 and, as a corollary, for its competitive binding with
chemokine ELC to said receptors. N-terminal polypeptide fragments of Acrp30R1L and Acrp30R1
30 comprising said motif non-productively bind to CCR7 and CCR10 and in so doing antagonize SLC
and ELC function through CCR7 and CCR10.

[0843] The instant invention is based on the discovery that N-terminal polypeptide fragments of the full-length Acrp30R1L or Acrp30R1 polypeptides comprising said motif have unexpected effects relating to antagonism of SLC and ELC chemokine function. These effects are unexpected and surprising given that full-length Acrp30R1L or Acrp30R1 polypeptide typically manifests no or 5 significantly reduced effects. To the extent that any effect is manifested by full-length Acrp30R1L or Acrp30R1 polypeptide, the levels of full-length polypeptide required for an effect render it unfeasible in most instances as a potential treatment for humans at this time. In contrast, N-terminal polypeptide fragments of full-length Acrp30R1L or Acrp30R1 polypeptides comprising said motif are radically more effective and thus can be provided at levels that are feasible for treatments in humans.

10 [0844] Thus, the invention is drawn to Acrp30R1L and Acrp30R1 polypeptide fragments, polynucleotides encoding said Acrp30R1L and Acrp30R1 polypeptide fragments, vectors comprising said Acrp30R1L and Acrp30R1 polynucleotides, and cells recombinant for said Acrp30R1L and Acrp30R1 polynucleotides, as well as to pharmaceutical and physiologically acceptable compositions comprising said Acrp30R1L and Acrp30R1 polypeptide fragments and methods of administering said 15 Acrp30R1L and Acrp30R1 pharmaceutical and physiologically acceptable compositions in order to reduce chronic inflammation, reduce atherosclerosis, reduce HIV replication in secondary lymphoid tissue, increase myelopoiesis, or treat SLC- or ELC-related diseases or disorders. Assays for identifying antagonists of said SLC- or ELC-related activity are also part of the invention.

[0845] In a first aspect, the invention features a purified, isolated, or recombinant Acrp30R1L 20 or Acrp30R1 polypeptide fragment that has significantly greater activity than a full-length Acrp30R1L or Acrp30R1 polypeptide, wherein said activity is selected from but not restricted to antagonism of SLC function or antagonism of ELC function. In preferred embodiments, Acrp30R1L polypeptide fragments having unexpected activity are selected from amino acids 21-46, 21-47, 21-48, 21-49, 21-50, 21-51, 21-52, 21-53, 21-54, 21-55, 21-56, 21-57, 21-58, 21-59, 21-60, 21-61, 21-62, 21- 25 63, 21-64, 21-65, 21-66, 21-67, 21-68, 21-69, 21-70, 21-71, 21-72, 21-73, 21-74, 21-75, 21-76, 21-77, 21-78, 21-79, 21-80, 21-81, 21-82, 21-83, 21-84, 21-85, 21-86, 21-87, 21-88, 21-89, 21-90, 21-91, 21- 92, 21-93, 21-94, 21-95, 21-96, 21-97, 21-98, 21-99, 21-100, 21-101, 21-102, 21-103, 21-104, 21-105, 21-106, 21-107, 21-108, 21-109, 21-110, 21-111, 21-112, 21-113, 21-114, 21-115, 21-116, 21-117, 21-118, 21-119, 21-120, 21-121, 21-122, 21-123, 21-124, 21-125, 21-126, 21-127, 21-128, 21-129, 30 21-130, 21-131, 21-132, or 21-133 of aforesaid Acrp30R1L amino acid sequence and wherein amino acid 21 is taken to be understood to denote the N-terminal amino acid of full-length mature Acrp30R1L polypeptide lacking the signal peptide.

[0846] In other preferred embodiments, Acrp30R1 polypeptide fragments having unexpected activity are selected from amino acids 21-46, 21-47, 21-48, 21-49, 21-50, 21-51, 21-52, 21-53, 21-54, 21-55, 21-56, 21-57, 21-58, 21-59, 21-60, 21-61, 21-62, 21-63, 21-64, 21-65, 21-66, 21-67, 21-68, 21-69, 21-70, 21-71, 21-72, 21-73, 21-74, 21-75, 21-76, 21-77, 21-78, 21-79, 21-80, 21-81, 21-82, 21-83, 5 21-84, 21-85, 21-86, or 21-87 of aforesaid Acrp30R1 amino acid sequence and wherein amino acid 21 is taken to be understood to denote the N-terminal amino acid of full-length mature Acrp30R1 polypeptide lacking the signal peptide.

[0847] In yet other preferred embodiments, Acrp30R1L polypeptide fragments having unexpected activity are selected from amino acids 21-46, 21-125, 21-126, 21-131, 21-132, and 21-133 10 of aforesaid Acrp30R1L amino acid sequence and wherein amino acid 21 is taken to be understood to denote the N-terminal amino acid of full-length mature Acrp30R1L polypeptide. In still other preferred embodiments, Acrp30R1 polypeptide fragments having unexpected activity are selected from amino acids 21-46 and 21-87 of aforesaid Acrp30R1 amino acid sequence and wherein amino acid 21 is taken to be understood to denote the N-terminal amino acid of full-length mature Acrp30R1 15 polypeptide.

[0848] In other further preferred embodiments, said polypeptide fragment comprises an amino acid sequence at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to the corresponding consecutive amino acids of aforesaid Acrp30R1L amino acid sequence or aforesaid Acrp30R1 amino acid sequence.

20 [0849] In other preferred embodiments, said Acrp30R1L polypeptide fragments are made by recombinant means or by proteolytic cleavage of full-length mature Acrp30R1L polypeptide. In other preferred embodiments, said Acrp30R1 polypeptide fragments are made by recombinant means or by proteolytic cleavage of full-length mature Acrp30R1 polypeptide lacking the signal peptide.

[0850] Particularly preferred proteolytic fragment of full-length mature Acrp30R1L 25 polypeptide lacking the signal peptide is about amino acids 21-46 made by collagenase cleavage of aforesaid Acrp30R1L amino acid sequence at about position 46 and wherein amino acid 21 is taken to be understood to denote the N-terminal amino acid of full-length mature Acrp30R1L polypeptide lacking the signal peptide. Other particularly preferred proteolytic fragment of full-length mature Acrp30R1L polypeptide lacking the signal peptide is about amino acids 21-46 made by MMP-1 30 cleavage of aforesaid Acrp30R1L amino acid sequence at about position 46 and wherein amino acid 21 is taken to be understood to denote the N-terminal amino acid of full-length mature Acrp30R1L polypeptide lacking the signal peptide. Other particularly preferred proteolytic fragment of full-length

mature Acrp30R1L polypeptide lacking the signal peptide is about amino acids 21-125 made by plasmin cleavage of aforesaid Acrp30R1L amino acid sequence at about position 125 and wherein amino acid 21 is taken to be understood to denote the N-terminal amino acid of full-length mature Acrp30R1L polypeptide lacking the signal peptide. Other particularly preferred proteolytic fragment 5 of full-length mature Acrp30R1L polypeptide lacking the signal peptide is about amino acids 21-126 made by plasmin cleavage of aforesaid Acrp30R1L amino acid sequence at about position 126 and wherein amino acid 21 is taken to be understood to denote the N-terminal amino acid of full-length mature Acrp30R1L polypeptide lacking the signal peptide. Other particularly preferred is proteolytic fragment of full-length mature Acrp30R1L polypeptide lacking the signal peptide is about amino acids 10 21-131 made by plasmin cleavage of aforesaid Acrp30R1L amino acid sequence at about position 131.

Other particularly preferred proteolytic fragment of full-length mature Acrp30R1L polypeptide lacking the signal peptide is about amino acids 21-132 made by plasmin cleavage of aforesaid Acrp30R1L amino acid sequence at about position 132 and wherein amino acid 21 is taken to be understood to denote the N-terminal amino acid of full-length mature Acrp30R1L polypeptide lacking the signal 15 peptide. Other particularly preferred proteolytic fragment of full-length mature Acrp30R1L polypeptide lacking the signal peptide is about amino acids 21-133 made by precerebellin processing protease cleavage of aforesaid Acrp30R1L amino acid sequence at about position 133 and wherein amino acid 21 is taken to be understood to denote the N-terminal amino acid of full-length mature Acrp30R1L polypeptide lacking the signal peptide.

20 [0851] Particularly preferred proteolytic fragment of full-length mature Acrp30R1 polypeptide lacking the signal peptide is about amino acids 21-46 made by collagenase cleavage of aforesaid Acrp30R1 amino acid sequence at about position 46 and wherein amino acid 21 is taken to be understood to denote the N-terminal amino acid of full-length mature Acrp30R1 polypeptide lacking the signal peptide. Other particularly preferred proteolytic fragment of full-length mature 25 Acrp30R1 polypeptide lacking the signal peptide is about amino acids 21-46 made by MMP-1 cleavage of aforesaid Acrp30R1 amino acid sequence at about position 46 and wherein amino acid 21 is taken to be understood to denote the N-terminal amino acid of full-length mature Acrp30R1 polypeptide lacking the signal peptide.

[0852] In a second aspect, the invention features a pharmaceutical or physiologically acceptable composition comprising, consisting essentially of, or consisting of, said Acrp30R1L or Acrp30R1 polypeptide fragment described in the first aspect and, alternatively, a pharmaceutical or physiologically acceptable diluent.

[0853] In a third aspect, the invention features a method of preventing or treating an immune-related disorder comprising providing or administering to an individual in need of such treatment said pharmaceutical or physiologically acceptable composition described in the second aspect. Preferably, said immune-related disorder is selected from but not restricted to allograft rejection or delayed type 5 hypersensitivity.

[0854] In a fourth aspect, the invention features a method of preventing or treating atherosclerosis comprising providing or administering to an individual in need of such treatment said pharmaceutical or physiologically acceptable composition described in the second aspect.

[0855] In a fifth aspect, the invention features a method of preventing or treating the 10 inflammation associated with ischemic neurodegeneration in brain tissue comprising providing or administering to an individual in need of such treatment said pharmaceutical or physiologically acceptable composition described in the second aspect.

[0856] In a sixth aspect, the invention features a method of preventing or treating the 15 inflammation associated with stroke comprising providing or administering to an individual in need of such treatment said pharmaceutical or physiologically acceptable composition described in the second aspect.

[0857] In a seventh aspect, the invention features a method of preventing or treating the 20 inflammation associated with myocardial infarction comprising providing or administering to an individual in need of such treatment said pharmaceutical or physiologically acceptable composition described in the second aspect.

[0858] In an eighth aspect, the invention features a method of preventing or treating the 25 inflammation associated with reperfusion injury comprising providing or administering to an individual in need of such treatment said pharmaceutical or physiologically acceptable composition described in the second aspect.

[0859] In a ninth aspect, the invention features a method of preventing or treating HIV 30 infection comprising providing or administering to an individual in need of such treatment said pharmaceutical or physiologically acceptable composition described in the second aspect.

[0860] In a tenth aspect, the invention features a method of preventing or treating an 35 inflammation-related disorder comprising providing or administering to an individual in need of such treatment said pharmaceutical or physiologically acceptable composition described in the second aspect. Preferably, said inflammation-related disorder is selected from but not restricted to rheumatoid

arthritis, inflammatory bowel disease, insulin dependent diabetes mellitus (Type 1 diabetes), systemic lupus erythematosus, psoriasis, allergic asthma, or septic shock.

[0861] In an eleventh aspect, the invention features a method of facilitating engraftment and expansion of infused hematopoietic stem cells or myeloid progenitor cells comprising providing or 5 administering to an individual in need of such treatment said pharmaceutical or physiologically acceptable composition described in the second aspect.

[0862] In a twelfth aspect, the invention features a method of facilitating engraftment and expansion within NOD/scid mice of infused human hematopoietic stem cells or myeloid progenitor cells related to the preconditioning regimen comprising a xenogeneic *in vivo* model of human 10 lymphoid and myeloid leukemia (Dialynas, DP et al., Blood 97:3218-25 (2001) which disclosure is hereby incorporated by reference in its entirety).

[0863] In a thirteenth aspect, the invention features a method of using said Acrp30R1L or Acrp30R1 polypeptide fragment described in the first aspect in *in vitro* cell migration assays as an internal control for intact CCR7 or CCR10 function.

15 [0864] The invention further features a method of using said Acrp30R1L or Acrp30R1 polypeptide fragment described in the first aspect in *in vitro* cell migration assays as an internal control for an intact T cell, B cell, or mature dendritic cell chemotactic response (Patel, DD et al., Clin Immunol 99:43-52 (2001) which disclosure is hereby incorporated by reference in its entirety).

[0865] In a fourteenth aspect, the invention provides for an antibody that specifically binds 20 said preferred polypeptide fragment of mature Acrp30R1L lacking the signal peptide but not to full-length mature Acrp30R1L polypeptide lacking the signal peptide. Further preferred is said antibody that recognizes a non-conformational or conformational epitope of said preferred polypeptide fragment of mature Acrp30R1L lacking the signal peptide. Further preferred is said antibody that neutralizes 25 the capacity of said preferred polypeptide fragment of mature Acrp30R1L lacking the signal peptide to antagonize SLC or ELC chemokine function.

[0866] Further preferred is a method wherein a mouse is immunized with said preferred polypeptide fragment of mature Acrp30R1L lacking the signal peptide. Further preferred is a method wherein monoclonal antibodies from said mouse are screened for binding to said preferred polypeptide fragment of mature Acrp30R1L lacking the signal peptide but not to full-length mature Acrp30R1L 30 polypeptide lacking the signal peptide. Further preferred is a method wherein monoclonal antibodies derived from said mouse are screened by enzyme-linked immunosorbent assay (ELISA) for binding to said preferred polypeptide fragment of mature Acrp30R1L but not to mature Acrp30R1L polypeptide

lacking the signal peptide. Further preferred is a method wherein said antibody is screened for the capacity to sterically or allosterically neutralize the capacity of said preferred polypeptide fragment of mature Acrp30R1L lacking the signal peptide to antagonize SLC or ELC chemokine function. Further preferred is a method of humanizing said monoclonal antibody. Methods of generating said 5 monoclonal antibody and of establishing specificity by methods including ELISA are well known to those skilled in the art. Methods of screening said antibody for neutralization of the capacity of said preferred polypeptide fragment of mature Acrp30R1L lacking the signal peptide to antagonize SLC or ELC chemokine function are well known to those skilled in the art and include, but are not limited to: contacting the antibody with said preferred polypeptide fragment of mature Acrp30R1L lacking the 10 signal peptide, incubation of said contacted Acrp30R1L polypeptide fragment with radiolabeled SLC or ELC in the presence of CCR7 or CCR10 chemokine receptor, and determination of whether said antibody contact blocks the capacity of said preferred polypeptide fragment of mature Acrp30R1L lacking the signal peptide to compete with SLC or ELC for binding to CCR7 or CCR10 (Gosling, J et al., *J Immunol* 164:2851-6 (2000) which disclosure is hereby incorporated by reference in its entirety).
15 Methods of humanizing said monoclonal antibody are well known to those skilled in the art.

[0867] In a fifteenth aspect, the invention provides for an antibody that specifically binds said preferred polypeptide fragment of Acrp30R1 but not full-length mature Acrp30R1 polypeptide lacking the signal peptide. Further preferred is said antibody that recognizes a non-conformational or conformational epitope of said preferred polypeptide fragment of Acrp30R1. Further preferred is said 20 antibody that neutralizes the capacity of said preferred Acrp30R1 polypeptide fragment to antagonize SLC or ELC chemokine function.

[0868] Further preferred is a method wherein a mouse is immunized with said preferred polypeptide fragment of mature Acrp30R1 lacking the signal peptide. Further preferred is a method wherein monoclonal antibodies from said mouse are screened for binding to said preferred polypeptide 25 fragment of mature Acrp30R1 lacking the signal peptide but not to full-length mature Acrp30R1 polypeptide lacking the signal peptide. Further preferred is a method wherein monoclonal antibodies derived from said mouse are screened by enzyme-linked immunosorbent assay (ELISA) for binding to said preferred polypeptide fragment of mature Acrp30R1 but not to mature Acrp30R1 polypeptide lacking the signal peptide. Further preferred is a method wherein said antibody is screened for the 30 capacity to sterically or allosterically neutralize the capacity of said preferred polypeptide fragment of mature Acrp30R1 lacking the signal peptide to antagonize SLC or ELC chemokine function. Further preferred is a method of humanizing said monoclonal antibody. Methods of generating said monoclonal antibody and of establishing specificity by methods including ELISA are well known to

those skilled in the art. Methods of screening said antibody for neutralization of the capacity of said preferred polypeptide fragment of mature Acrp30R1 lacking the signal peptide to antagonize SLC or ELC chemokine function are well known to those skilled in the art and include, but are not limited to: contacting the antibody with said preferred polypeptide fragment of mature Acrp30R1 lacking the signal peptide, incubation of said contacted Acrp30R1 polypeptide fragment with radiolabeled SLC or ELC in the presence of CCR7 or CCR10 chemokine receptor, and determination of whether said antibody contact blocks the capacity of said preferred polypeptide fragment of mature Acrp30R1 lacking the signal peptide to compete with SLC or ELC for binding to CCR7 or CCR10 (Gosling, J et al., *J Immunol* 164:2851-6 (2000) which disclosure is hereby incorporated by reference in its entirety).

10 Methods of humanizing said monoclonal antibody are well known to those skilled in the art.

[0869] In a sixteenth aspect, the invention features a pharmaceutical or physiologically acceptable composition comprising, consisting essentially of, or consisting of, said antibody of the fourteenth aspect directed to said preferred Acrp30R1L polypeptide fragment and, alternatively, a pharmaceutical or physiologically acceptable diluent.

15 [0870] In a seventeenth aspect, the invention features a pharmaceutical or physiologically acceptable composition comprising, consisting essentially of, or consisting of, said antibody of the fifteenth aspect directed to said preferred Acrp30R1 polypeptide fragment and, alternatively, a pharmaceutical or physiologically acceptable diluent.

[0871] In an eighteenth aspect, the invention features a method of using said antibody specific for said preferred Acrp30R1L or Acrp30R1 polypeptide fragment described in the fourteenth or sixteenth aspect in a method of measuring the amount of said Acrp30R1L or Acrp30R1 polypeptide fragment in a clinical sample for diagnosis of immune dysfunction. Preferably, said clinical sample is selected from the group consisting of blood, serum, plasma, urine, and saliva. Preferably, said immune dysfunction is selected from but not restricted to recurrent infection, innate immunodeficiency, or 25 acquired immunodeficiency. Elevated levels of said Acrp30R1L or Acrp30R1 polypeptide fragment are expected to be contributory to said immune dysfunction.

[0872] In a nineteenth aspect, the invention features a method of using said antibody specific for said preferred Acrp30R1L or Acrp30R1 polypeptide fragment described in the fourteenth or sixteenth aspect in a method of measuring the amount of said Acrp30R1L or Acrp30R1 polypeptide fragment in a clinical sample for cancer stratification. Preferably, said clinical sample is selected from but not restricted to blood, serum, plasma, urine, and saliva. Preferably, said cancer is selected from but not restricted to melanoma, squamous cell carcinoma of the skin, breast carcinoma, lung small-cell

carcinoma, colon carcinoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma, prostatic carcinoma, pancreatic carcinoma, osteosarcoma, uterine carcinoma, ovarian carcinoma, chondrosarcoma, endometrial cancer, testicular carcinoma, renal carcinoma, hepatic carcinoma, lung non-small-cell carcinoma, T cell acute lymphoblastic leukemia, B cell acute lymphoblastic leukemia, acute myeloid leukemia, chronic lymphocytic leukemia, or multiple myeloma. Elevated levels of said Acrp30R1L or Acrp30R1 are expected to be contributory to an impaired anti-tumor response.

[0873] In a twentieth aspect, the invention features a method of treating cancer characterized by diagnostically determined elevated level of said preferred Acrp30R1L or Acrp30R1 polypeptide fragment comprising providing or administering to an individual in need of such treatment said pharmaceutical or physiologically acceptable composition described in the sixteenth or seventeenth aspect wherein the antibody comprising said composition neutralizes the capacity of said preferred Acrp30R1L or Acrp30R1 polypeptide fragment to antagonize SLC or ELC chemokine function. Said neutralization is expected to facilitate the *in vivo* anti-tumor response. Preferably, said cancer is selected from but not restricted to melanoma, squamous cell carcinoma of the skin, breast carcinoma, lung small-cell carcinoma, colon carcinoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma, prostatic carcinoma, pancreatic carcinoma, osteosarcoma, uterine carcinoma, ovarian carcinoma, chondrosarcoma, endometrial cancer, testicular carcinoma, renal carcinoma, hepatic carcinoma, lung non-small-cell carcinoma, T cell acute lymphoblastic leukemia, B cell acute lymphoblastic leukemia, acute myeloid leukemia, chronic lymphocytic leukemia, or multiple myeloma.

[0874] In a twenty-first aspect, the invention provides for a non-antibody compound that specifically binds said preferred polypeptide fragment of mature Acrp30R1L lacking the signal peptide but not to full-length mature Acrp30R1L polypeptide lacking the signal peptide. Preferred said compound is selected from but not restricted to small molecular weight organic or inorganic compound, protein, peptide, carbohydrate, or lipid. Further preferred is said antibody that recognizes a non-conformational or conformational epitope of said preferred polypeptide fragment of mature Acrp30R1L lacking the signal peptide. Further preferred is said antibody that neutralizes the capacity of said preferred polypeptide fragment of mature Acrp30R1L lacking the signal peptide to antagonize SLC or ELC chemokine function.

[0875] Further preferred are methods of screening for one or more compounds capable of neutralizing the capacity of said preferred polypeptide fragment of mature Acrp30R1L lacking the signal peptide to antagonize SLC or ELC chemokine function. Said methods are well known to those skilled in the art and include, but are not limited to: contacting the compound with said preferred polypeptide fragment of mature Acrp30R1L lacking the signal peptide, incubation of said contacted

Acp30R1L polypeptide fragment with radiolabeled SLC or ELC in the presence of CCR7 or CCR10 chemokine receptor, and determination of whether said contact with compound blocks the capacity of said preferred polypeptide fragment of mature Acp30R1L lacking the signal peptide to compete with SLC or ELC for binding to CCR7 or CCR10 (Gosling, J et al., J Immunol 164:2851-6 (2000) which disclosure is hereby incorporated by reference in its entirety).

[0876] In a twenty-second aspect, the invention provides for a non-antibody compound that specifically binds said preferred polypeptide fragment of mature Acp30R1 lacking the signal peptide but not to full-length mature Acp30R1 polypeptide lacking the signal peptide. Preferred said compound is selected from but not restricted to small molecular weight organic or inorganic compound, protein, peptide, carbohydrate, or lipid. Further preferred is said antibody that recognizes a non-conformational or conformational epitope of said preferred polypeptide fragment of mature Acp30R1 lacking the signal peptide. Further preferred is said antibody that neutralizes the capacity of said preferred polypeptide fragment of mature Acp30R1 lacking the signal peptide to antagonize SLC or ELC chemokine function.

[0877] Further preferred are methods of screening for one or more compounds capable of neutralizing the capacity of said preferred polypeptide fragment of mature Acp30R1 lacking the signal peptide to antagonize SLC or ELC chemokine function. Said methods are well known to those skilled in the art and include, but are not limited to: contacting the compound with said preferred polypeptide fragment of mature Acp30R1 lacking the signal peptide, incubation of said contacted Acp30R1L polypeptide fragment with radiolabeled SLC or ELC in the presence of CCR7 or CCR10 chemokine receptor, and determination of whether said contact with compound blocks the capacity of said preferred polypeptide fragment of mature Acp30R1L lacking the signal peptide to compete with SLC or ELC for binding to CCR7 or CCR10 (Gosling, J et al., J Immunol 164:2851-6 (2000) which disclosure is hereby incorporated by reference in its entirety).

[0878] In a twenty-third aspect, the invention features a pharmaceutical or physiologically acceptable composition comprising, consisting essentially of, or consisting of, said compound of the twenty-first aspect directed to said preferred Acp30R1L polypeptide fragment and, alternatively, a pharmaceutical or physiologically acceptable diluent.

[0879] In a twenty-fourth aspect, the invention features a pharmaceutical or physiologically acceptable composition comprising, consisting essentially of, or consisting of, said antibody of the twenty-second aspect directed to said preferred Acp30R1 polypeptide fragment and, alternatively, a pharmaceutical or physiologically acceptable diluent.

[0880] In a twenty-fifth aspect, the invention features a method of treating cancer characterized by diagnostically determined elevated level of said preferred Acrp30R1L or Acrp30R1 polypeptide fragment comprising providing or administering to an individual in need of such treatment said pharmaceutical or physiologically acceptable composition described in the twenty-third or twenty-fourth aspect wherein the compound comprising said composition neutralizes the capacity of said preferred Acrp30R1L or Acrp30R1 polypeptide fragment to antagonize SLC or ELC chemokine function. Said neutralization is expected to facilitate the *in vivo* anti-tumor response. Preferably, said cancer is selected from but not restricted to melanoma, squamous cell carcinoma of the skin, breast carcinoma, lung small-cell carcinoma, colon carcinoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma, prostatic carcinoma, pancreatic carcinoma, osteosarcoma, uterine carcinoma, ovarian carcinoma, chondrosarcoma, endometrial cancer, testicular carcinoma, renal carcinoma, hepatic carcinoma, lung non-small-cell carcinoma, T cell acute lymphoblastic leukemia, B cell acute lymphoblastic leukemia, acute myeloid leukemia, chronic lymphocytic leukemia, or multiple myeloma.

15 CRPF Polypeptides CRPF-1, CRPF-2A, CRPF-2B, CRPF-2C, and CRPF-3

[0881] CRPF polypeptide CRPF-1 comprises the amino acid sequence: SHHH. CRPF polypeptide CRPF-2A comprises the amino acid sequence: AANSKVAFSAVRSTNH. CRPF polypeptide CRPF-2B comprises the amino acid sequence: AANSKVAFSAVR. CRPF polypeptide CRPF-2C comprises the amino acid sequence: STNH. CRPF polypeptide CRPF-3 comprises the amino acid sequence: SGSAKVAFSATRSTNH. Also preferred are polynucleotides encoding the CRPF polypeptides.

[0882] The CRPF polypeptides are related to cerebellin. The CRPF polypeptides have unexpected and novel function as described below.

[0883] Cerebellin is a 16 amino acid polypeptide liberated by specific proteolytic cleavage 25 from a precursor polypeptide, precerebellin (Urade, Y et al., Proc Natl Acad Sci USA 88:1069-1073 (1991) which disclosure is hereby incorporated by reference in its entirety). Cerebellin is conserved in sequence from human to chicken. Cerebellin is highly expressed in cerebellum, weakly expressed in other areas of the brain, and possibly also expressed in adrenomedullary tissue (Satoh, F et al., Journal of Endocrinology 154:27-34 (1997) which disclosure is hereby incorporated by reference in its 30 entirety). Cerebellin is also expressed by the tumor tissues of adrenal tumor, ganglioneuroblastoma, and neuroblastoma (Satoh, F et al., Journal of Endocrinology 154:27-34 (1997) which disclosure is hereby incorporated by reference in its entirety). Cerebellin elicits epinephrine and norepinephrine

release by human adrenomedullary cells (Albertin, G et al., *Neuropeptides* 34:7-11 (2000) which disclosure is hereby incorporated by reference in its entirety). Norepinephrine has been shown to up-regulate expression of the proinflammatory cytokine interferon gamma by the Th1 subset of CD4+ T lymphocytes (Swanson, MA et al., *J Immunol* 166:232-240 (2001) which disclosure is hereby incorporated by reference in its entirety). Interferon gamma plays an important role in the development of cytolytic T lymphocytes (CTL). Recently, a precerebellin-like protein was shown to be part of the acute phase response in rainbow trout (Gerwick, L et al., *Developmental and Comparative Immunology* 24:597-600 (2000) which disclosure is hereby incorporated by reference in its entirety).

10 [0884] The instant invention is based on the discovery that the CRPF polypeptides oppose the proinflammatory action of cerebellin.

[0885] Thus, the invention is drawn to CRPF polypeptides, polynucleotides encoding said CRPF polypeptides, vectors comprising said CRPF polynucleotides, and cells recombinant for said CRPF polynucleotides, as well as to pharmaceutical and physiologically acceptable compositions 15 comprising said CRPF polypeptides and methods of administering said CRPF pharmaceutical and physiologically acceptable compositions in order to reduce inflammation or to treat inflammation-related disorders. Assays for identifying antagonists of CRPF polypeptide function are also part of the invention.

[0886] In a first aspect, the invention features purified, isolated, artificially synthesized, or 20 recombinant CRPF polypeptide that has anti-inflammatory activity. In preferred embodiment, CRPF polypeptide is CRPF-1 polypeptide comprising the amino acid sequence: SHHH. In other preferred embodiment, CRPF polypeptide is CRPF-2A polypeptide comprising the amino acid sequence: AANSKVAFSAVRSTNH. In other preferred embodiment, CRPF polypeptide is CRPF-2B polypeptide comprising the amino acid sequence AANSKVAFSAVR. In other preferred embodiment, 25 CRPF polypeptide is CRPF-2C polypeptide comprising the amino acid sequence STNH. In other preferred embodiment, CRPF polypeptide is CRPF-3 comprising the amino acid sequence SGSAKVAFSATRSTNH.

[0887] In a second aspect, the invention features a pharmaceutical or physiologically acceptable composition comprising, consisting essentially of, or consisting of said CRPF polypeptides 30 described in the first aspect and, alternatively, a pharmaceutical or physiologically acceptable diluent.

[0888] In a third aspect, the invention features a method of reducing inflammation comprising providing or administering to individuals in need of reducing inflammation said pharmaceutical or

physiologically acceptable composition described in the second aspect. In preferred embodiment, said administration of said pharmaceutical or physiologically acceptable composition is systemic or local.

[0889] In a fourth aspect, the invention features a method of preventing or treating psoriasis comprising providing or administering to an individual in need of such treatment said pharmaceutical or physiologically acceptable composition described in the second aspect.
5

[0890] In a fifth aspect, the invention features a method of preventing or treating atherosclerosis comprising providing or administering to an individual in need of such treatment said pharmaceutical or physiologically acceptable composition described in the second aspect.

[0891] In a sixth aspect, the invention features a method of preventing or treating the 10 inflammation associated with ischemic neurodegeneration in brain tissue comprising providing or administering to an individual in need of such treatment said pharmaceutical or physiologically acceptable composition described in the second aspect.

[0892] In a seventh aspect, the invention features a method of preventing or treating the 15 inflammation associated with stroke comprising providing or administering to an individual in need of such treatment said pharmaceutical or physiologically acceptable composition described in the second aspect.

[0893] In an eighth aspect, the invention features a method of preventing or treating the 20 inflammation associated with myocardial infarction comprising providing or administering to an individual in need of such treatment said pharmaceutical or physiologically acceptable composition described in the second aspect.

[0894] In a ninth aspect, the invention features a method of preventing or treating the inflammation associated with reperfusion injury comprising providing or administering to an individual in need of such treatment said pharmaceutical or physiologically acceptable composition described in the second aspect.

25 [0895] In a tenth aspect, the invention features a method of preventing or treating the inflammation associated with rheumatoid arthritis comprising providing or administering to an individual in need of such treatment said pharmaceutical or physiologically acceptable composition described in the second aspect.

[0896] In an eleventh aspect, the invention features a method of preventing or treating the 30 inflammation associated with inflammatory bowel disease comprising providing or administering to an

individual in need of such treatment said pharmaceutical or physiologically acceptable composition described in the second aspect.

5 [0897] In a twelfth aspect, the invention features a method of preventing or treating the inflammation associated with systemic lupus erythematosus comprising providing or administering to an individual in need of such treatment said pharmaceutical or physiologically acceptable composition described in the second aspect.

10 [0898] In a thirteenth aspect, the invention features a method of preventing or treating the inflammation associated with allergic asthma comprising providing or administering to an individual in need of such treatment said pharmaceutical or physiologically acceptable composition described in the second aspect.

[0899] In a fourteenth aspect, the invention features a method of preventing or treating the inflammation associated with insulin dependent diabetes mellitus (Type 1 diabetes) comprising providing or administering to an individual in need of such treatment said pharmaceutical or physiologically acceptable composition described in the second aspect.

15 [0900] In a fifteenth aspect, the invention features a method of preventing or treating the inflammation associated with septic shock comprising providing or administering to an individual in need of such treatment said pharmaceutical or physiologically acceptable composition described in the second aspect.

20 [0901] In a sixteenth aspect, the invention features a method of preventing or treating the inflammation associated with multiple sclerosis comprising providing or administering to an individual in need of such treatment said pharmaceutical or physiologically acceptable composition described in the second aspect.

25 [0902] In a seventeenth aspect, the invention features a method of preventing or treating the inflammation associated with elevated cerebellin level in blood or locally at a site of inflammation or other disorder described herein comprising providing or administering to an individual in need of such treatment said pharmaceutical or physiologically acceptable composition described in the second aspect. Preferably the magnitude of said elevated cerebellin level is at least 10%, 20%, 30%, 35%, 40%, 50%, 75%, 100%, or 500%. Preferably, cerebellin level is elevated at the site where said CRPF polypeptide is administered.

30 [0903] In an eighteenth aspect, the invention features a method of suppressing an *in vivo* Th1 immune response in an animal model of inflammopathology or immune dysfunction comprising

providing or administering to said animal said pharmaceutical or physiologically acceptable composition described in the second aspect. Preferably said animal is a mammal. Most preferably, said animal is a rodent.

[0904] In a nineteenth aspect, the invention provides for an antibody that specifically binds to 5 said CRPF polypeptide described in the first aspect. Further preferred is said antibody that neutralizes the capacity of said CRPF polypeptide to reduce inflammation. Further preferred is said antibody that neutralizes the capacity of said CRPF polypeptide to bind to a ligand. Further preferred is said antibody that neutralizes the capacity of said CRPF polypeptide to bind to its receptor.

[0905] In a twentieth aspect, the invention features a pharmaceutical or physiologically 10 acceptable composition comprising, consisting essentially of, or consisting of, said neutralizing antibody described in the nineteenth aspect specific for said CRPF polypeptide described in the first aspect and, alternatively, a pharmaceutical or physiologically acceptable diluent.

[0906] In a twenty-first aspect, the invention features a method of using said antibody 15 described in the nineteenth aspect specific for said CRPF polypeptide described in the first aspect to measure the amount of said CRPF polypeptide in a clinical sample in order to diagnose or stratify an inflammation-related disorder. Preferably, said clinical sample is selected from the group consisting of blood, serum, plasma, urine, saliva, and lymphoid tissue. Preferably, said inflammation-related disorder is selected from but not restricted to those described herein. Reduced levels of said CRPF polypeptide are expected to be contributory to said inflammation-related disorder.

20 [0907] In a twenty-second aspect, the invention features a method of using said antibody described in the nineteenth aspect specific for said CRPF polypeptide described in the first aspect to measure the amount of said CRPF polypeptide in a clinical sample obtained from a cancer patient for the purpose of characterizing said cancer. Preferably, said clinical sample is selected from but not restricted to blood, serum, plasma, urine, saliva, and cancer tissue. Preferably, said method is selected 25 from but not restricted to ELISA or immunohistochemistry. Preferably, said cancer is selected from but not restricted to adrenal tumor, ganglioneuroblastoma, neuroblastoma, melanoma, squamous cell carcinoma of the skin, breast carcinoma, lung small-cell carcinoma, colon carcinoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma, prostatic carcinoma, pancreatic carcinoma, osteosarcoma, uterine carcinoma, ovarian carcinoma, chondrosarcoma, endometrial cancer, testicular carcinoma, 30 renal carcinoma, hepatic carcinoma, lung non-small-cell carcinoma, T cell acute lymphoblastic leukemia, B cell acute lymphoblastic leukemia, acute myeloid leukemia, chronic lymphocytic

leukemia, or multiple myeloma. Elevated levels of said CRPF polypeptide (for example released by the tumor cells) are expected to be contributory to an impaired anti-tumor response.

[0908] In a twenty-third aspect, the invention features a method of treating cancer characterized by elevated level of said CRPF polypeptide described in the first aspect comprising 5 providing or administering to an individual in need of such treatment said pharmaceutical or physiologically acceptable composition described in the twentieth aspect. Neutralization of the capacity of said CRPF polypeptide suppress an inflammatory response is expected to facilitate the *in vivo* anti-tumor response. Preferably, said cancer is selected from but not restricted to adrenal tumor, ganglioneuroblastoma, neuroblastoma, melanoma, squamous cell carcinoma of the skin, breast 10 carcinoma, lung small-cell carcinoma, colon carcinoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma, prostatic carcinoma, pancreatic carcinoma, osteosarcoma, uterine carcinoma, ovarian carcinoma, chondrosarcoma, endometrial cancer, testicular carcinoma, renal carcinoma, hepatic carcinoma, lung non-small-cell carcinoma, T cell acute lymphoblastic leukemia, B cell acute lymphoblastic leukemia, acute myeloid leukemia, chronic lymphocytic leukemia, or multiple 15 myeloma.

[0909] In a twenty-fourth aspect, the invention provides for a non-antibody compound that specifically binds said CRPF polypeptide described in the first aspect but not to unrelated polypeptide. Preferred said compound is selected from but not restricted to small molecular weight organic or inorganic compound, protein, peptide, carbohydrate, or lipid. Further preferred is said compound that 20 neutralizes the capacity of said CRPF polypeptide to reduce inflammation.

[0910] Further preferred are methods of screening for one or more compounds capable of neutralizing the capacity of said CRPF polypeptide fragment to reduce inflammation. Said methods are well known to those skilled in the art and include, but are not limited to: administering to a mouse model of inflammopathology said CRPF polypeptide in the presence of said antibody; administering to 25 said mouse model said CRPF polypeptide in the absence of said antibody; and determination of whether said antibody abrogates the capacity of said CRPF polypeptide to reduce inflammation in said mouse model.

[0911] In a twenty-fifth aspect, the invention features a pharmaceutical or physiologically acceptable composition comprising, consisting essentially of, or consisting of, said neutralizing 30 compound of the twenty-fourth aspect and, alternatively, a pharmaceutical or physiologically acceptable diluent.

[0912] In a twenty-sixth aspect, the invention features a method of treating a cancer patient having elevated levels of said CRPF polypeptide described in the first aspect comprising providing or administering to an individual in need of such treatment said pharmaceutical or physiologically acceptable composition described in the twenty-fifth aspect. Preferably, methods for measuring the amount of said CRPF polypeptide fragment include but are not restricted to said methods described in the twenty-second aspect. Neutralization of the capacity of said CRPF polypeptide to suppress an inflammatory response is expected to facilitate the *in vivo* anti-tumor response. Preferably, said cancer is selected from but not restricted to adrenal tumor, ganglioneuroblastoma, neuroblastoma, melanoma, squamous cell carcinoma of the skin, breast carcinoma, lung small-cell carcinoma, colon carcinoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma, prostatic carcinoma, pancreatic carcinoma, osteosarcoma, uterine carcinoma, ovarian carcinoma, chondrosarcoma, endometrial cancer, testicular carcinoma, renal carcinoma, hepatic carcinoma, lung non-small-cell carcinoma, T cell acute lymphoblastic leukemia, B cell acute lymphoblastic leukemia, acute myeloid leukemia, chronic lymphocytic leukemia, or multiple myeloma.

15

Uses of antibodies

[0913] Antibodies of the present invention have uses that include, but are not limited to, methods known in the art to purify, detect, and target the polypeptides of the present invention including both *in vitro* and *in vivo* diagnostic and therapeutic methods. For example, the antibodies have use in immunoassays for qualitatively and quantitatively measuring levels of antigen-bearing substances, including the polypeptides of the present invention, in biological samples (See, e.g., Harlow *et al.*, 1988). The antibodies may also be used in therapeutic compositions for killing cells expressing the protein or reducing the levels of the protein in the body.

[0914] The invention further relates to antibodies that act as agonists or antagonists of the polypeptides of the present invention. For example, the present invention includes antibodies that disrupt the receptor/ligand interactions with the polypeptides of the invention either partially or fully. These antibodies may act as agonists for either all or less than all of the biological activities affected by ligand-mediated receptor activation. The antibodies may be specified as agonists or antagonists for biological activities comprising specific activities disclosed herein. The above antibody agonists can be made using methods known in the art. See e.g., WO 96/40281; US Patent 5,811,097; Deng *et al.*, 20 (1998) Blood. 92(6):1981-1988; Chen *et al.*, (1998), Cancer Res. 58(16):3668-3678; Harrop *et al.*, (1998), J. Immunol. 161(4):1786-1794; Zhu, *et al.* (1998), Cancer Res. 58(15):3209-3214; Yoon, *et al.* (1998), J. Immunol. 160(7):3170-3179; Prat *et al.*, (1998), J. Cell. Sci. 111(Pt2):237-247; Pitard *et al.*, (1997), J. Immunol. Methods. 205(2):177-190; Liautard *et al.*, (1997), Cytokine. 9(4):233-241;

Carlson *et al.*, (1997), J. Biol. Chem. 272(17):11295-11301; Taryman, *et al.*, (1995), Neuron. 14(4):755-762; Muller *et al.*, (1998), Structure. 6(9):1153-1167; Bartunek *et al.*, (1996), Cytokine. 8(1):14-20.

[0915] As discussed above, antibodies of the polypeptides of the invention can, in turn, be utilized to generate anti-idiotypic antibodies that "mimic" polypeptides of the invention using techniques well known to those skilled in the art (see, e.g. Greenspan and Bona (1989), FASEB J. 7(5):437-444 and Nissinoff, (1991), J. Immunol. 147(8): 2429-2438). For example, antibodies which bind to and competitively inhibit polypeptide multimerization or binding of a polypeptide of the invention to ligand can be used to generate anti-idiotypes that "mimic" the polypeptide multimerization or binding domain and, as a consequence, bind to and neutralize polypeptide or its ligand. Such neutralization anti-idiotypic antibodies can be used to bind a polypeptide of the invention or to bind its ligands/receptors, and thereby block its biological activity.

Immunoaffinity Chromatography

[0916] Antibodies prepared as described herein are coupled to a support. Preferably, the antibodies are monoclonal antibodies, but polyclonal antibodies may also be used. The support may be any of those typically employed in immunoaffinity chromatography, and the antibodies may be coupled to the support using any standard reagent. After coupling the antibody to the support, the support is contacted with a sample which contains a target polypeptide whose isolation, purification or enrichment is desired.

[0917] Thereafter, the support is washed with an appropriate wash solution, and the specifically bound target polypeptide is eluted from the support using the high pH or low pH elution solutions typically employed in immunoaffinity chromatography.

25

Expression Of Genset Gene Products

Evaluation of Expression Levels and Patterns of GENSET polypeptide-encoding mRNAs

[0918] The spatial and temporal expression patterns of GENSET polypeptide-encoding mRNAs, as well as their expression levels, may be determined using any suitable method.

[0919] In one embodiment, expression levels and patterns of GENSET polypeptide-encoding mRNAs is analyzed by solution hybridization with probes (see, e.g., WO 97/05277). Briefly, an RNA complementary to the mRNA of interest is labeled and derivatized with a capturable moiety, e.g., biotin. After hybridization in solution with mRNA isolated from cells or tissues of interest, 5 unhybridized probe is removed by digestion, and the remaining, hybridized RNA is captured, e.g., on a microtitration plate, and quantified.

[0920] In another embodiment, the GENSET polypeptide-encoding cDNAs, or fragments thereof, may also be tagged with nucleotide sequences for the serial analysis of gene expression (SAGE) (see, e.g., UK Patent Application No. 2 305 241 A).

10 [0921] Quantitative analysis of GENSET gene expression may also be performed using arrays. For example, quantitative analysis of gene expression may be performed with GENSET polynucleotides, or fragments thereof in a complementary DNA microarray as described by Schena *et al.* (1995) Science 270:467-470 and Schena *et al.* (1996), PNAS, 93(20):10614-10619, Pietu *et al.*, (1996) Genome Research 6:492-503) or using any other microarray technology. Briefly, GENSET 15 polypeptide-encoding cDNAs or fragments thereof are arrayed onto slides, and the arrays are hybridized with probes derived from mRNA of cells or tissues of interest. Following washing, the arrays are scanned using a fluorescence laser scanning device fitted with a custom filter set. Accurate differential expression measurements are obtained by taking the average of the ratios of two independent hybridizations.

20 [0922] Alternatively, expression analysis of GENSET genes can be done through high density nucleotide arrays as described by Lockhart *et al.*, (1996) Nature Biotechnology 14: 1675-1680 and Sosnowski, *et al.*, (1997) PNAS 94:1119-1123. Oligonucleotides of 15-50 nucleotides corresponding to sequences of a GENSET polynucleotide or fragments thereof are synthesized directly on the chip or synthesized and then addressed to the chip. Labeled cDNA probes are synthesized from the 25 appropriate mRNA population and then randomly fragmented to an average size of 50 to 100 nucleotides. The probes are then hybridized to the chip. After washing, the label is detected and quantified. Comparative analysis of the intensity of the signal originating from cDNA probes on the same target oligonucleotide in different cDNA samples indicates a differential expression of the GENSET polypeptide-encoding mRNA.

Uses of GENSET gene expression data

[0923] Once the expression levels and patterns of a GENSET polypeptide-encoding mRNA has been determined, this information may be used to design GENSET gene specific markers for detection, identification, screening and diagnostic purposes as well as to design DNA constructs with
5 an expression pattern similar to a GENSET gene expression pattern.

Detection of GENSET polypeptide expression and/or biological activity

[0924] The invention further relates to methods of detection of GENSET polypeptide expression and/or biological activity in a biological sample using the polynucleotide and polypeptide sequences described herein. Such methods can be used, for example, as a screen for normal or abnormal GENSET polypeptide expression and/or biological activity and, thus, can be used diagnostically.
10

Detection of GENSET polypeptides

[0925] The invention further relates to methods of detection of GENSET polypeptides in a cell or sample. In certain embodiments, the presence of polypeptides in a cell or sample is detected indirectly, by detecting the presence of mRNA encoding the polypeptide. For example, a labeled polynucleotide probe can be used in a method to detect a GENSET polypeptide-encoding mRNA, wherein the presence of the mRNA is indicative of the expression of the GENSET polypeptide-
20 encoding gene.

[0926] Consequently, the invention comprises methods for detecting the presence of a polynucleotide of the invention in a sample, the method comprising bringing into contact said sample and a nucleic acid probe or a plurality of nucleic acid probes which hybridize to the polynucleotide, and detecting the hybrid complex formed between said probe or said plurality of probes and said polynucleotide. In certain embodiments, the probe or probes are labeled or are immobilized on a substrate.
25

[0927] In another embodiment, the polynucleotide is detected using an amplification reaction, wherein the sample is contacted with amplification reaction reagents, an amplification reaction is performed to synthesize amplification products containing said region of said selected nucleotide sequence; and the amplification products are detected. In a preferred embodiment, when the
30

polynucleotide to be amplified is a RNA molecule, preliminary reverse transcription and synthesis of a second cDNA strand are first performed in order to provide a DNA template to be amplified.

[0928] Alternatively, a method of detecting GENSET polypeptide expression in a test sample can be accomplished using any product which binds to a GENSET polypeptide of the present 5 invention or portion thereof. Such products may be antibodies, binding fragments of antibodies, polypeptides able to bind specifically to GENSET polypeptides or fragments thereof, including GENSET polypeptide agonists and antagonists. Detection of specific binding to the antibody indicates the presence of a GENSET polypeptide in the sample (e.g., ELISA).

[0929] Consequently, the invention is also directed to a method for detecting specifically the 10 presence of a GENSET polypeptide in a biological sample, said method comprising bringing into contact the biological sample with a product able to bind to a polypeptide of the invention or fragments thereof; allowing the product to bind to the polypeptide to form a complex; and detecting the complex. In a preferred embodiment, the product is an antibody, e.g., an antibody that is immobilized on a substrate.

[0930] The present invention also relates to kits that can be used in the detection of GENSET 15 polypeptide-encoding gene expression products, e.g. containing a compound that specifically binds a GENSET polypeptide (e.g. binding proteins, antibodies or binding fragments thereof (e.g. F(ab')² fragments) or a GENSET polypeptide-encoding mRNA (e.g. a complementary probe or primer), disposed within a container means. The kit can further comprise ancillary reagents, including buffers 20 and the like.

Detection of GENSET polypeptide biological activity

[0931] The invention further includes methods of detecting specifically a GENSET 25 polypeptide biological activity, and to identify compounds capable of modulating the activity of a GENSET polypeptide. Assessing the GENSET polypeptide biological activity may be performed by the detection of a change in any cellular property associated with the GENSET polypeptide, using a variety of techniques, including those described herein. To identify modulators of the polypeptides, a control is preferably used. For example, a control sample includes all of the same reagents but lacks the compound or agent being assessed; it is treated in the same manner as the test sample.

[0932] The present invention also relates to kits that can be used in the detection of GENSET 30 polypeptide biological activity, e.g., including substrates for GENSET polypeptides, GENSET-binding

compounds, antibodies to GENSET polypeptides, etc., disposed within a container means. The kit can further comprise ancillary reagents, including buffers and the like.

Identification of a specific context of GENSET polypeptide-encoding gene expression

- 5 [0933] When the expression pattern of a GENSET polypeptide-encoding mRNA shows that a GENSET polypeptide-encoding gene is specifically expressed in a given context, probes and primers specific for this gene as well as antibodies binding to the GENSET polypeptide-encoding polynucleotide may then be used as markers for the specific context. Examples of specific contexts are: specific expression in a given tissue/cell or tissue/cell type, expression at a given stage of
10 development of a process such as embryo development or disease development, expression in response to a particular compound or drug, or specific expression in a given organelle. Such primers, probes, and antibodies are useful commercially to identify tissues/cells/organelles of unknown origin, for example, forensic samples, differentiated tumor tissue that has metastasized to foreign bodily sites, or to differentiate different tissue types in a tissue cross-section.
- 15 [0934] Determination of tissue/cell/organelle identity is based on methods that detect the presence or absence of the mRNA (or corresponding cDNA or protein) in a tissue/cell sample using methods well known in the art (e.g., hybridization, PCR based methods, immunoassays, immunochemistry, ELISA). Therefore, the invention encompasses uses of the polynucleotides, polypeptides, and antibodies of the invention as tissue markers.

20

Identification of Tissue Types or Cell Species by Means of Labeled Tissue Specific Antibodies

- [0935] Identification of specific tissues is accomplished by the visualization of tissue specific antigens by means of antibody preparations which are conjugated, directly (e.g., green fluorescent protein) or indirectly to a detectable marker. Selected labeled antibody species bind to their specific
25 antigen binding partner in tissue sections, cell suspensions, or in extracts of soluble proteins from a tissue sample to provide a pattern for qualitative or semi-qualitative interpretation.

A. Immunohistochemical Techniques

- [0936] Purified, high-titer antibodies, prepared as described above, are conjugated to a
30 detectable marker, as described, for example, by Fudenberg, (1980) Chap. 26 in: Basic 503 Clinical

Immunology, 3rd Ed. Lange, Los Altos, California or Rose *et al.*, (1980) Chap. 12 in: Methods in Immunodiagnosis, 2d Ed. John Wiley Sons, New York.

[0937] Antibodies can be labeled using any suitable method, e.g. fluorescent labels, enzymes (e.g. horseradish peroxidase), ferritin (for detection using EM), or radiolabeling. In one embodiment, 5 cryostat sections of the unknown tissue and known control are mounted and each slide covered with different dilutions of the antibody preparation. Following incubation, excess fluid is blotted away, and the marker detected. The antigen found in the tissues by the above procedure can be quantified by measuring the intensity of color or fluorescence on the tissue section, and calibrating that signal using appropriate standards.

10

B. Identification of Tissue Specific Soluble Proteins

[0938] The visualization of tissue specific proteins and identification of unknown tissues from that procedure is carried out using the labeled antibody reagents and detection strategy as described for immunohistochemistry; however the sample is prepared according to an electrophoretic technique to 15 distribute the proteins extracted from the tissue in an orderly array on the basis of molecular weight for detection, e.g. by western blotting. See, e.g., Davis *et al.*, Basic Methods in Molecular Biology, ed., Elsevier Press, NY (1986), Section 19-3.

Screening and diagnosis of abnormal GENSET polypeptide expression and/or biological activity

20 [0939] Moreover, antibodies and/or primers or probes specific for GENSET polypeptide expression may also be used to identify abnormal GENSET polypeptide expression and/or biological activity, and subsequently to screen and/or diagnose disorders associated with abnormal GENSET polypeptide expression. For example, a particular disease may result from lack of expression, over-expression, or under-expression of a GENSET polypeptide-encoding mRNA. By comparing mRNA 25 expression patterns and quantities in samples taken from healthy individuals with those from individuals suffering from a particular disorder, genes responsible for this disorder may be identified.

Screening for specific disorders

[0940] The present invention also relates to methods and uses of GENSET polypeptides for 30 identifying individuals having elevated or reduced levels of GENSET polypeptides, which individuals

are likely to benefit from therapies to suppress or enhance GENSET polypeptide-encoding gene expression, respectively. One example of such methods and uses comprises detecting the presence in a biological sample of a GENSET polypeptide-encoding gene product (mRNA or protein); comparing the amount of the GENSET polypeptide-encoding gene product present in said sample with that of a 5 control sample; and determining whether the sample has a reduced or elevated level of GENSET gene expression compared to the control sample.

[0941] A biological sample from a subject affected by, or at risk of developing, any disease or condition associated with a GENSET polypeptide can be screened for the presence of increased or decreased levels of GENSET gene product, relative to a normal population (standard or control), with 10 an increased or decreased level of the GENSET polypeptide relative to the normal population being indicative of predisposition to or a present indication of the disease or condition, or any symptom associated with the disease or condition. Such individuals would be candidates for therapies, e.g., treatment with pharmaceutical compositions comprising the GENSET polypeptide, a polynucleotide encoding the GENSET polypeptide, or any other compound that affects the expression or activity of 15 the GENSET polypeptide. Generally, the identification of elevated levels of the GENSET polypeptide in a patient would be indicative of an individual that would benefit from treatment with agents that suppress GENSET polypeptide expression or activity, and the identification of low levels of the GENSET polypeptide in a patient would be indicative of an individual that would benefit from agents that induce GENSET expression or activity.

[0942] Biological samples suitable for use in this method include any biological fluids, including, but not limited to, blood, saliva, milk, and urine, as well as tissue samples such as biopsies. Cell cultures or cell extracts derived, for example, from tissue biopsies can also be used. In preferred embodiments, the biological sample is taken from animals presenting any symptom associated with any disease or condition associated with a GENSET gene product. In accordance with this method, 25 the presence in the sample of altered (e.g. increased or decreased) levels of the GENSET product indicates that the subject is predisposed to the disease or condition.

[0943] The diagnostic methodologies described herein are applicable to both humans and non-human mammals.

30 Detection of GENSET gene mutations

[0944] The invention also encompasses methods and uses of GENSET polynucleotides to detect mutations in GENSET polynucleotides of the invention. When the mutation was proven to be

associated with a disease, the detection of such mutations may be used for screening and diagnosis purposes.

[0945] In one embodiment, an oligonucleotide probe matrix may advantageously be used to detect mutations occurring in GENSET genes and preferably in their regulatory regions. For this 5 particular purpose, probes are specifically designed to have a nucleotide sequence allowing their hybridization to the genes that carry known mutations (either by deletion, insertion or substitution of one or several nucleotides). By known mutations, it is meant, mutations on the GENSET genes that have been identified according, for example to the technique used by Huang *et al.*, (1996) Cancer Res 56(5):1137-1141 or Samson *et al.*, (1996) Nature, 382(6593):722-725.

10 [0946] Another technique that is used to detect mutations in GENSET genes is the use of a high-density DNA array. Each oligonucleotide probe constituting a unit element of the high density DNA array is designed to match a specific subsequence of a GENSET genomic DNA or cDNA. Thus, an array consisting of oligonucleotides complementary to subsequences of the target gene sequence is used to determine the identity of the target sequence with the wild gene sequence, measure its amount, 15 and detect differences between the target sequence and the reference wild gene sequence of the GENSET gene. In one such embodiment, a 4L tiled array is used (see, e.g., Chee *et al.*, (1996) Science. 274:610-614).

Construction of DNA constructs with a GENSET gene expression pattern

20 [0947] In addition, characterization of the spatial and temporal expression patterns and expression levels of GENSET polypeptide-encoding mRNAs is also useful for constructing expression vectors capable of producing a desired level of gene product in a desired spatial or temporal manner, as discussed below.

25 DNA Constructs That Direct Temporal And Spatial GENSET Gene Expression In Recombinant Cell Hosts And In Transgenic Animals.

[0948] In order to study the physiological and phenotypic consequences of a lack of synthesis of a GENSET polypeptide, both at the cellular level and at the multi-cellular organism level, the invention also encompasses DNA constructs and recombinant vectors enabling a conditional 30 expression of a specific allele of a GENSET polypeptide-encoding genomic sequence or cDNA,

including alleles that include substitutions, deletions, or additions of one or more bases within the sequence.

[0949] In one embodiment, a DNA construct is used that is based on the tetracycline resistance operon *tet* from *E. coli* transposon Tn10 for controlling the GENSET gene expression, such as described by Gossen *et al.*, (1992) PNAS 89:5547-5551; Gossen *et al.*, (1995) Science 268:1766-1769; and Furth P.A. *et al.* (1994) PNAS 91:9302-9306.

[0950] In another embodiment, a DNA construct is used that comprises, from 5'-end to 3'-end: a first nucleotide sequence that is found in the GENSET polypeptide-encoding genomic sequence; a nucleotide sequence comprising a positive selection marker (e.g. *neo*); and a second nucleotide sequence that is found in the GENSET polypeptide-encoding genomic sequence, and is located downstream of the first GENSET nucleotide sequence. In a preferred embodiment, the construct also comprises a negative selection marker (e.g. thymidine kinase, hygromycin beta, hprt, Diphtheria toxin A fragment) located upstream of the first GENSET nucleotide sequence or downstream from the second GENSET nucleotide sequence (see, e.g., Thomas *et al.* (1986), Cell. 44:419-428; Te Riele *et al.* (1990), Nature. 348:649-651; Van der Lugt *et al.* (1991), Gene 105:263-267; Reid *et al.*, (1990) PNAS 87:4299-4303; Nada *et al.*, (1993) Cell 73:1125-1135; Yagi, T., *et al.* (1990), PNAS 87:9918-9922; Thomas *et al.* (1986; 1987), Mansour *et al.*(1988) and Koller *et al.*, (1992) Annu. Rev. Immunol. 10:705-730).

[0951] In another embodiment, vectors are used that involve the use of the Cre-loxP system (Hoess *et al.*, (1986) Nucleic Acids Res. 14:2287-2300). The Cre-loxP system used in combination with a homologous recombination technique has been described by Gu H. *et al.*, (1993) Cell 73:1155-1164 and Gu H. *et al.*, (1994) Science 265:103-106. Briefly, a nucleotide sequence of interest to be inserted in a targeted location of the genome harbors at least two *loxP* sites in the same orientation and located at the respective ends of a nucleotide sequence to be excised from the recombinant genome. The excision event requires the presence of the recombinase (Cre) enzyme within the nucleus of the recombinant cell host, which may be supplied in any of a number of ways, e.g., by incubating the recombinant cell hosts in a culture medium containing this enzyme, by injecting the Cre enzyme directly into the desired cell, by lipofection of the enzyme into the cells, or by otherwise transfected the cell with a Cre coding sequence under the control of an appropriate promoter (see, e.g., Baubonis *et al* (1993) Nucleic Acids Res. 21(9):2025-9); Araki *et al.*, (1995) PNAS 92(1):160-4; Gu *et al.* (1993); Sauer *et al.*, (1988) PNAS 85:5166-5170; Gu *et al.*(1994); Zou, *et al.*, (1994) Curr. Biol. 4:1099-1103 ; Anton and Graham, (1995), J. Virol., 69: 4600-4606; and Kanegae *et al.*, (1995) Nucl. Acids Res. 23:3816-3821).

[0952] The DNA constructs described above may be used to introduce a desired nucleotide sequence of the invention, preferably a GENSET genomic sequence or a GENSET cDNA sequence, and most preferably an altered copy of a GENSET genomic or cDNA sequence, within a predetermined location of the targeted genome, leading either to the generation of an altered copy of a targeted gene (knock-out homologous recombination) or to the replacement of a copy of the targeted gene by another copy sufficiently homologous to allow an homologous recombination event to occur (knock-in homologous recombination).

Modifying genset polypeptide expression and/or biological activity

[0953] Modifying endogenous GENSET expression and/or biological activity is expressly contemplated by the present invention.

Screening for compounds that modulate GENSET expression and/or biological activity

[0954] The present invention further relates to compounds able to modulate GENSET expression and/or biological activity and methods to use these compounds. Such compounds may interact with the regulatory sequences of GENSET genes or they may interact with GENSET polypeptides directly or indirectly.

Compounds Interacting With GENSET Regulatory Sequences

[0955] The present invention also concerns a method for screening substances or molecules that are able to interact with the regulatory sequences of a GENSET gene, such as for example promoter or enhancer sequences in untranscribed regions of the genomic DNA, as determined using any techniques known to those skilled in the art, or such as regulatory sequences located in untranslated regions of GENSET mRNA.

[0956] Sequences within untranscribed or untranslated regions of polynucleotides of the invention may be identified by comparison to databases containing known regulatory sequence such as transcription start sites, transcription factor binding sites, promoter sequences, enhancer sequences, 5'UTR and 3'UTR elements (Pesole *et al.*, (2000) Nucleic Acids Res, 28(1):193-196; <http://igs-server.cnrs-mrs.fr/~gauthere/UTR/index.html>). Alternatively, the regulatory sequences of interest may be identified through conventional mutagenesis or deletion analyses of reporter plasmids.

[0957] Following the identification of potential GENSET regulatory sequences, proteins which interact with these regulatory sequences may be identified as described below.

[0958] Any of a number of methods can be used to identify molecules capable of interacting with the regulatory sequence of a GENSET gene, such as gel retardation assays (see, e.g., Fried and Crothers, (1981) Nucleic Acids Res. 9:6505-6525; Garner and Revzin, (1981) Nucleic Acids Res 9:3047-3060 ; and Dent and Latchman (1993) The DNA mobility shift assay. In: Transcription Factors: A Practical Approach (Latchman DS, ed.) pp: 1-26, Oxford: IRL Press). Nucleic acids encoding proteins which are able to interact with the promoter sequence of a GENSET gene may also be identified by using a one-hybrid system (e.g., the Matchmaker One-Hybrid System kit from Clontech (Catalog Ref. no. K1603-1)).

Ligands interacting with GENSET polypeptides

[0959] For the purpose of the present invention, a ligand means a molecule, such as a protein, a peptide, an antibody or any synthetic chemical compound capable of binding to a GENSET protein or one of its fragments or variants or to modulate the expression of the polynucleotide coding for GENSET or a fragment or variant thereof.

[0960] In one embodiment, a biological sample or a defined molecule (e.g. a molecule generated through combinatorial chemistry) to be tested as a putative ligand of a GENSET protein is brought into contact with the purified GENSET protein, in order to determine if a complex is formed between this protein and a component of the sample or the defined molecule. The interaction between such molecules and the protein can be assessed in any way, e.g., using microdialysis coupled to HPLC, using affinity capillary electrophoresis, etc. (see, e.g., Wang, *et al.* (1997), Chromatographia, 44 : 205-208; Bush *et al.*, (1997), J. Chromatogr., 777 : 311-328).

[0961] Any type of compound can be tested, using any method, for interaction with a GENSET polypeptide or polynucleotide in the present methods, including, but not limited to, peptides, drugs, fatty acids, lipoproteins, or small molecules, and may be obtained from any source. For example, the molecule to be tested is labeled with a detectable label, such as a fluorescent, radioactive, or enzymatic tag and placed in contact with immobilized GENSET protein, or a fragment thereof under conditions which permit specific binding to occur. After removal of non-specifically bound molecules, bound molecules are detected using appropriate means.

A. Candidate ligands obtained from random peptide libraries

[0962] In a particular embodiment of the screening method, the putative ligand is the expression product of a DNA insert contained in a phage vector, e.g. from a random peptide phage library comprising peptides of from 8 to 20 amino acids in length (Parmley and Smith (1988) *Gene* 73:305-318; Oldenburg *et al.* (1992) *PNAS* 89:5393-5397; Valadon *et al.* (1996) *J. Mol. Biol.*, 261:11-22; Lucas (1994) In: Development and Clinical Uses of Haemophilus b Conjugate; Westerink (1995) *PNAS* 92:4021-4025; Felici (1991) *J. Mol. Biol.*, 222:301-310).

[0963] Once the ligand library in recombinant phages has been constructed, the phage population is brought into contact with the immobilized GENSET protein, and, following washing, the phages that bind specifically to the GENSET protein are either eluted by a buffer (acid pH) or immunoprecipitated using an antibody specific to the GENSET protein. The isolated phage is subsequently amplified by an over-infection of bacteria (for example E. coli). The selection step may be repeated several times, preferably 2-4 times, in order to select the more specific recombinant phage clones.

B. Candidate ligands obtained by competition experiments

[0964] Alternatively, peptides, drugs or small molecules which bind to polypeptides of the present invention may be identified in competition experiments. In one such assay, a GENSET protein, or a fragment thereof, is immobilized to a surface, such as a plastic plate. Increasing amounts of the peptides, drugs or small molecules are placed in contact with the immobilized GENSET protein, or a fragment thereof, in the presence of a detectable labeled known GENSET protein ligand. For example, the GENSET ligand may be detectably labeled with a fluorescent, radioactive, or enzymatic tag. The ability of the test molecule to bind the GENSET protein, or a fragment thereof, is determined by measuring the amount of detectably labeled known ligand bound in the presence of the test molecule. A decrease in the amount of known ligand bound to the GENSET protein, or a fragment thereof, when the test molecule is present indicated that the test molecule is able to bind to the GENSET protein, or a fragment thereof.

C. Candidate ligands obtained by affinity chromatography

[0965] Proteins or other molecules interacting with a polypeptide of the present invention can also be found using affinity columns which contain the GENSET protein, or a fragment thereof. The GENSET protein, or a fragment thereof, may be attached to the column using conventional techniques including chemical coupling to a suitable column matrix (e.g. agarose, Affi Gel®, etc.). In some embodiments of this method, the affinity column contains chimeric proteins in which the GENSET protein, or a fragment thereof, is fused to glutathion S transferase (GST). A mixture of cellular

proteins or pool of expressed proteins as described above is applied to the affinity column. Proteins or other molecules interacting with the GENSET protein, or a fragment thereof, attached to the column can then be isolated and analyzed, e.g., on 2-D electrophoresis gel as described in Ramunsen *et al.*, (1997), Electrophoresis, 18: 588-598. Alternatively, the proteins retained on the affinity column can 5 be purified by electrophoresis-based methods and sequenced. The same method can be used to isolate antibodies, to screen phage display products, or to screen phage display human antibodies.

D. Candidate ligands obtained by optical biosensor methods

[0966] Proteins interacting with a polypeptide of the present invention, can also be screened by using an Optical Biosensor (see, e.g., Edwards and Leatherbarrow (1997) Anal. Bioch. 246:1-6; 10 Szabo *et al.*, (1995) Curr Opin Struct Biol 5, 699-705. This technique permits the detection of interactions between molecules in real time, without the need of labeled molecules. This technique is based on the surface plasmon resonance (SPR) phenomenon. In one embodiment, a GENSET polypeptide, or fragment thereof, is attached to a surface (such as a carboxymethyl dextran matrix) comprising one side of a cell through which flows the candidate molecule to be assayed. A light beam 15 is directed towards the side of the surface that does not contain the sample to be tested and is reflected by said surface. The binding of a candidate ligand molecules cause a change in the refraction index on the surface, which change is detected as a change in the SPR signal. This technique may also be performed by immobilizing eukaryotic or prokaryotic cells or lipid vesicles exhibiting an endogenous or a recombinantly expressed GENSET protein at their surface.

20

E. Candidate ligands obtained through a two-hybrid screening assay

[0967] The yeast two-hybrid system is designed to study protein-protein interactions *in vivo* (Fields and Song, 1989; US Patent Nos. US 5,667,973 and 5,283,173), and relies upon the fusion of a bait protein to the DNA binding domain of the yeast Gal4 protein. The general procedure of library 25 screening by the two-hybrid assay may be performed as described by Harper *et al.*, (1993), *Cell* 75:805-816; Cho *et al.*, (1998) *PNAS* 95(7):3752-3757; Fromont-Racine *et al.*, (1997) *Nature Genetics* 16(3):277-282. In typical embodiments, the bait protein comprises a polypeptide of the present invention, and the "prey" comprises a human cDNA library constructed such that the human cDNA insert is fused to a nucleotide sequence in the vector that encodes the transcriptional domain of the 30 GAL4 protein. In another embodiment, interaction between the GENSET polypeptide or a fragment or variant thereof with cellular proteins is assessed using the Matchmaker Two Hybrid System 2 (Catalog No. K1604-1, Clontech).

Compounds Modulating GENSET biological activity

[0968] Another method of screening for compounds that modulate GENSET expression and/or biological activity is by measuring the effects of test compounds on specific biological activity, e.g. a GENSET biological activity in a host cell or in an *in vitro* assay. A GENSET biological activity can include any of the activities, functions, or properties described herein for any GENSET polypeptide or polynucleotide. In one embodiment, a nucleic acid construct encoding a GENSET polypeptide is introduced into a host cell, and the host cell is maintained under conditions appropriate for expression of the encoded GENSET polypeptide, whereby the nucleic acid is expressed. The host cell is then contacted with a test agent, wherein a detection of a change in any GENSET polypeptide-associated property in the presence of the agent indicates that the agent alters GENSET biological activity. In a particular embodiment, the invention relates to a method of identifying an agent which is an activator of GENSET biological activity, wherein detection of an increase of any GENSET polypeptide-associated property in the presence of the agent indicates that the agent activates GENSET biological activity. In another particular embodiment, the invention relates to a method of identifying an agent which is an inhibitor of GENSET biological activity, wherein detection of a decrease of any GENSET polypeptide-associated property in the presence of the agent indicates that the agent inhibits GENSET biological activity. In another particular embodiment, a high throughput screen is used to identify agents that activate (enhance) or inhibit GENSET biological activity (see, e.g., WO 98/45438).

Methods of Screening for Compounds Modulating GENSET Gene Expression and/or Activity

[0969] The present invention also relates to methods of screening compounds for their ability to modulate (e.g. increase or inhibit) the activity or expression of GENSET polypeptides and polynucleotides. More specifically, the present invention relates to methods of testing compounds for their ability either to increase or to decrease expression or activity of GENSET polypeptides and polynucleotides. The assays are performed *in vitro* or *in vivo*.

In vitro methods

[0970] *In vitro*, cells expressing GENSET polypeptides, or capable of expressing GENSET polypeptides, are incubated in the presence and absence of the test compound. By determining the level of GENSET expression in the presence of the test compound or the level of GENSET biological activity in the presence of the test compound, compounds can be identified that suppress or enhance GENSET expression or activity. Alternatively, constructs comprising a GENSET regulatory sequence

operably linked to a reporter gene (e.g. luciferase, chloramphenicol acetyl transferase, LacZ, green fluorescent protein, etc.) can be introduced into host cells and the effect of the test compounds on expression of the reporter gene detected. Consequently, the present invention encompasses a method for screening molecules that modulate the expression of a GENSET gene, said screening method comprising cultivating a prokaryotic or a eukaryotic cell that has been transfected with a nucleotide sequence encoding either a GENSET polypeptide, placed under the control of its own promoter, or a detectable polypeptide, placed under the control of a GENSET 5' regulatory region; bringing into contact the cultivated cell with a molecule to be tested; and quantifying the expression of the GENSET or detectable polypeptide in the presence of the molecule. The method can also be performed using fragments, variants, or derivatives of any of the GENSET polypeptides or 5' regulatory regions.

[0971] The quantification of the expression of a GENSET polypeptide may be realized either at the mRNA level (using for example Northen blots, RT-PCR, preferably quantitative RT-PCR with primers and probes specific for the GENSET mRNA of interest) or at the protein level (using polyclonal or monoclonal antibodies in immunoassays such as ELISA or RIA assays, Western blots, or immunochemistry).

[0972] In a further embodiment, the GENSET 5' regulatory region includes the 5'UTR region of a GENSET polynucleotide sequence, and/or a promoter sequence which is endogenous or exogenous with respect to the GENSET 5'UTR sequence.

[0973] The invention further relates to a method for the production of a pharmaceutical composition comprising identifying a molecule that modulates the expression of a GENSET gene using any of the herein-described methods, and combining the identified molecule with a physiologically acceptable carrier.

[0974] Kits for the screening candidate substances for the ability to modulate the expression of a GENSET gene. Preferably, such kits comprise a recombinant vector comprising a GENSET 5' regulatory region or a regulatory active fragment or a variant thereof, operably linked to a polynucleotide encoding a detectable protein or a GENSET protein or a fragment or a variant thereof.

[0975] Another object of the present invention comprises methods and kits for the screening of candidate substances that interact with a GENSET polypeptide, fragments or variants thereof. By their capacity to bind covalently or non-covalently to a GENSET protein, fragments or variants thereof, these substances or molecules may be advantageously used both *in vitro* and *in vivo*.

[0976] *In vitro*, said interacting molecules may be used as detection means in order to identify the presence of a GENSET protein in a sample, preferably a biological sample.

[0977] In one embodiment, a method is provided for the screening of a candidate substance, the method comprising providing a GENSET protein; bringing into contact the protein with the candidate substance; and determining whether a complex is formed between the polypeptide or fragment and the candidate substance.

5 [0978] The invention further relates to a method for the production of a pharmaceutical composition comprising identifying a substance that interacts with a GENSET polypeptide using any of the herein-described methods, fragments or variants thereof and furthermore mixing the identified substance with a physiologically acceptable carrier.

10 [0980] The invention further concerns a kit for the screening of a candidate substance interacting with the GENSET polypeptide, wherein said kit comprises a GENSET polypeptide, and optionally means to detect a complex formed between the polypeptide and the candidate substance. In one embodiment, the detection means comprises a monoclonal or polyclonal antibody binding to said GENSET protein or fragment or variant thereof.

15 In vivo methods

[0981] Compounds that suppress or enhance GENSET gene expression can also be identified using *in vivo* screens. In a typical assay, a test compound is administered (e.g. intravenously, intraperitoneally, intramuscularly, orally, or otherwise) to an animal, at a variety of dose levels, and the effect of the compound on GENSET gene expression is determined by comparing the levels of the 20 mRNA or protein encoded by the gene in tissues known to express the gene of interest, e.g., using Northern blots, immunoassays, PCR, etc.. Suitable test animals include, but are not limited to, rodents (e.g., mice and rats), primates, and rabbits. Humanized mice can also be used, that is mice in which the endogenous mouse protein is ablated (knocked out) and the homologous human protein introduced using standard transgenic approaches. Such mice thus express only the human form of a protein. 25 Humanized mice expressing only the human GENSET polypeptide can be used to study *in vivo* responses to potential agents regulating GENSET protein or mRNA levels. Such transgenic animals are useful for dissecting the biochemical and physiological steps of disease, and for development of therapies for disease intervention (see, e.g., Loring, *et al*, 1996).

[0982] In addition, the detection of any change in any GENSET gene-associated behavior or 30 characteristic of an animal following the administration of a compound can also be used as an indication that the compound modulates the expression or activity of the gene.

Uses for compounds modulating GENSET expression and/or biological activity

[0983] Using *in vivo* (or *in vitro*) systems, it may be possible to identify compounds that exert a tissue specific effect, for example, that increase GENSET expression or activity in one or more 5 particular tissues of interest, such as the adrenal gland, bone marrow, brain, cerebellum, colon, fetal brain, fetal kidney, fetal liver, heart, hypertrophic prostate, kidney, liver, lung, lymph ganglia, lymphocytes, muscle, ovary, pancreas, pituitary gland, placenta, prostate, salivary gland, spinal cord, spleen, stomach, intestine, substantia nigra, testis, thyroid, umbilical cord, or uterus. Screening procedures such as those described above are also useful for identifying agents for their potential use 10 in pharmacological intervention strategies. Agents that enhance GENSET gene expression or stimulate its activity may thus be used to induce any phenotype associated with a GENSET gene, or to treat disorders resulting from a deficiency of a GENSET polypeptide activity or expression. Compounds that suppress GENSET polypeptide expression or inhibit its activity can be used to treat any disease or condition associated with increased or deleterious GENSET polypeptide activity or 15 expression.

[0984] Also encompassed by the present invention is an agent which interacts with a GENSET gene or polypeptide directly or indirectly, and inhibits or enhances GENSET polypeptide expression and/or function. In one embodiment, the agent is an inhibitor which interferes with a GENSET polypeptide directly (e.g., by binding the GENSET polypeptide) or indirectly (e.g., by 20 blocking the ability of the GENSET polypeptide to have a GENSET biological activity). In a particular embodiment, an inhibitor of a GENSET protein is an antibody specific for the GENSET protein or a functional portion of the GENSET protein. Alternatively, the inhibitor can be an agent other than an antibody (e.g., small organic molecule, protein or peptide) which binds the GENSET polypeptide and blocks its activity. For example, the inhibitor can be an agent which mimics the 25 GENSET polypeptide structurally, but lacks its function (e.g. a dominant negative form of the protein). Alternatively, it can be an agent which binds to or interacts with a molecule which the GENSET polypeptide normally binds to or interacts with, thus blocking the GENSET polypeptide from doing so and preventing it from exerting the effects it would normally exert.

[0985] In another embodiment, the agent is an enhancer (activator) of a GENSET polypeptide 30 which increases the activity of the GENSET polypeptide (increases the effect of a given amount or level of GENSET polypeptide), increases the length of time it is effective (by preventing its degradation or otherwise prolonging the time during which it is active) or both either directly or

indirectly. For example, GENSET polynucleotides and polypeptides can be used to identify drugs which increase or decrease the ability of GENSET polypeptides to induce GENSET biological activity, which drugs are useful for the treatment or prevention of any disease or condition associated with a GENSET biological activity.

5 [0986] Thus the present invention relates to a method of inhibiting (partially or completely) a GENSET biological activity in a mammal (e.g., a human), the method comprising administering to the mammal an effective amount of an inhibitor of a GENSET polypeptide or polynucleotide. The invention also relates to a method of enhancing a GENSET biological activity in a mammal, the method comprising administering to the mammal an effective amount of an enhancer of a GENSET
10 polypeptide or polynucleotide.

Inhibiting GENSET gene expression

[0987] GENSET gene expression can be inhibited, e.g., for therapeutic applications, in any of a large number of ways, including by using an antisense tool or a triple helix tool that inhibits the
15 expression of the corresponding GENSET gene.

Antisense Approach

[0988] In antisense approaches, DNA and/or RNA sequences complementary to an mRNA are hybridized to the mRNA intracellularly, thereby blocking the expression of the protein encoded by the mRNA. Preferred methods for using antisense polynucleotide according to the present invention
20 are the procedures described by Sczakiel *et al.*, (1995) Trends Microbiol. 3(6):213-217; Green *et al.*, (1986) Ann. Rev. Biochem. 55:569-597 ; Izant and Weintraub, (1984) Cell 36(4):1007-15; Liu *et al.* (1994) PNAS 91: 4528-4262 ; Eckner *et al.*, (1991) EMBO J. 10:3513-3522).

[0989] Preferably, the antisense tools are chosen among the polynucleotides (15-200 bp long) that are complementary to GENSET mRNA, more preferably to the 5'end (e.g. the translation
25 initiation codon ATG) or to a splicing donor or acceptor site of the GENSET mRNA. In another embodiment, a combination of different antisense polynucleotides complementary to different parts of the desired targeted gene are used. The antisense molecules can be prepared in any way, including by reversing the orientation of a coding region of a GENSET gene in a cell, or by synthesizing an oligonucleotide in vitro and administering it to the cell.

30 [0990] Any antisense sequence complementary to any portion of any of the herein-described polynucleotides can be used, and can involve any number of modifications to the backbone, linkages,

or bases, a large number of which are known in the art. Suitable modifications, other forms of antisense molecules, and the preparation thereof, are taught, *inter alia*, in The Concise Encyclopedia Of Polymer Science And Engineering, pages 858-859, Kroschwitz, J. I., ed. John Wiley & Sons, 1990; Englisch et al., Angewandte Chemie, International Edition (1991), 30, 613; Sanghvi, et al., eds., 5 (1993) Antisense Research and Applications, CRC Press, Boca Raton; U.S. Patent 6,242,590; WO94/23026, WO 96/31523; WO 92/18522; European Patent Application No. 0 572 287 A2; WO 92/19732; Letsinger et al., PNAS (1989) 86: 6553-6556; Manoharan et al., Bioorg. Med. Chem. Let. (1994) 4:1053-1060; Manoharan et al., Ann. N.Y. Acad. Sci. (1992) 660:306-309; Manoharan et al., Bioorg. Med. Chem. Let. (1993) 3:2765-2770; Oberhauser et al., Nucl. Acids Res. (1992) 20:533-538; 10 Saison-Behmoaras et al., EMBO J. (1991) 10:1111-1118; Kabanov et al., FEBS Lett. (1990) 259: 327-330; Svinarchuk et al., Biochimie (1993) 75:49-54; Manoharan et al., Tetrahedron Lett. (1995) 36, 3651-3654; Shea et al., Nucl. Acids Res. (1990) 18:3777-3783; Manoharan et al., Nucleosides & Nucleotides (1995) 14: 969-973; Manoharan et al., Tetrahedron Lett. (1995) 36:3651-3654; Mishra et al., Biochim. Biophys. Acta (1995) 1264:229-237; Crooke et al., J. Pharmacol. Exp. Ther. (1996) 15 277:923-937; U.S. Patent 6,242,590.

[0991] Further included in the present invention is a method of high throughput screening of antisense nucleic acids and modified versions thereof for binding to targeted GENSET polynucleotide sequences or fragments thereof (see, e.g., U.S. Patent 6,022,691).

[0992] The appropriate level of antisense nucleic acids required to inhibit gene expression 20 may be determined using *in vitro* expression analysis. Typically, antisense molecules are introduced onto cell samples at a number of different concentrations preferably between 1×10^{-10} M to 1×10^{-4} M. Once the minimum concentration that can adequately control gene expression is identified, the optimized dose is translated into a dosage suitable for use *in vivo* or *ex vivo*. For example, an inhibiting concentration in culture of 1×10^{-7} translates into a dose of approximately 0.6 mg/kg 25 bodyweight.

[0993] An alternative to the antisense technology that is used according to the present invention comprises using ribozymes that will bind to a target sequence via their complementary polynucleotide tail and that will cleave the corresponding RNA by hydrolyzing its target site (see, e.g., Rossi et al., (1991) Pharmacol. Ther. 50:245-254 and Sczakiel et al. (1995)).

30 Triple Helix Approach

[0994] The GENSET genomic DNA may also be used to inhibit the expression of the GENSET gene based on intracellular triple helix formation (see, e.g., Griffin et al., (1989) Science

245:967-971). To carry out gene therapy strategies using the triple helix approach, the sequences of the GENSET genomic DNA are typically first scanned to identify 10-mer to 20-mer homopyrimidine or homopurine stretches which could be used in triple-helix based strategies for inhibiting GENSET gene expression. Following identification of candidate homopyrimidine or homopurine stretches, their 5 efficiency in inhibiting GENSET gene expression is assessed by introducing varying amounts of oligonucleotides containing the candidate sequences into tissue culture cells which express the GENSET gene. The oligonucleotides can be introduced into the cells using a variety of methods known to those skilled in the art, and treated cells are monitored for altered cell function or reduced GENSET gene expression.

10 [0995] The oligonucleotides which are effective in inhibiting gene expression in tissue culture cells may then be introduced *in vivo* using the techniques and at a dosage calculated based on the *in vitro* results, as described in the section entitled "Antisense Approach".

Treating GENSET gene-related disorders

[0996] The present invention further relates to methods, uses of GENSET polypeptides and polynucleotides, and uses of modulators of GENSET polypeptides and polynucleotides, for treating diseases/disorders associated with GENSET genes by increasing or decreasing GENSET gene activity and/or expression. These methodologies can be effected using compounds selected using screening protocols such as those described herein and/or by using the gene therapy and antisense approaches described in the art and herein. Gene therapy can be used to effect targeted expression of GENSET genes in any tissue, e.g. a tissue associated with the disease or condition to be treated. The GENSET coding sequence can be cloned into an appropriate expression vector and targeted to a particular cell type(s) to achieve efficient, high level expression. Introduction of the GENSET coding sequence into target cells can be achieved, for example, using particle mediated DNA delivery, (Haynes *et al.*, (1996) J Biotechnol. 44(1-3):37-42 and Maurer *et al.*, (1999) Mol Membr Biol. 16(1):129-40), direct injection of naked DNA (Levy *et al.*, (1996) Gene Ther. 3(3):201-11; and Felgner (1996) Hum Gene Ther. 7(15):1791-3), or viral vector mediated transport (Smith *et al.*, (1996) Antiviral Res. 32(2):99-115, Stone *et al.*, (2000) J Endocrinol. 164(2):103-18; Wu and Ataa (2000), Curr Opin Biotechnol. 11(2):205-8). Tissue specific effects can be achieved, for example, in the case of virus mediated transport by using viral vectors that are tissue specific, or by the use of promoters that are tissue specific. For instance, any tissue-specific promoter may be used to achieve specific expression, for example albumin promoters (liver specific; Pinkert *et al.*, 1987 Genes Dev. 1:268-277), lymphoid specific promoters (Calame *et al.*, 1988 Adv. Immunol. 43:235-275), promoters of T-cell receptors (Winoto *et al.*, 1989 EMBO J. 8:729-733) and immunoglobulins (Banerji *et al.*, 1983 Cell 33:729-740; Queen and Baltimore 1983 Cell 33:741-748), neuron-specific promoters (e.g. the neurofilament promoter; Byrne *et al.*, 1989 PNAS 86:5473-5477), pancreas-specific promoters (Edlunch *et al.*, 1985 Science 230:912-916) or mammary gland-specific promoters (milk whey promoter, U.S. Pat. No. 25 4,873,316 and European Application Publication No. 264, 166). Developmentally-regulated promoters can also be used, such as the murine homeobox promoters (Kessel *et al.*, 1990 Science 249:374-379) or the alpha-fetoprotein promoter (Campes *et al.*, 1989 Genes Dev. 3:537-546).

Combinatorial approaches can also be used to ensure that the GENSET coding sequence is activated in the target tissue (Butt and Karathanasis (1995) Gene Expr. 4(6):319-36; Miller and Whelan, (1997) Hum Gene Ther. 8(7):803-15). Antisense oligonucleotides complementary to GENSET mRNA can be used to selectively diminish or ablate the expression of the protein (Wagner, *et al.* (1996), Nat Biotechnol. 14(7):840-4)), for example, at sites of inflammation. More specifically, antisense constructs or antisense oligonucleotides can be used to inhibit the production of GENSET in

high expressing cells, e.g., by transfecting target cells with an expression vector comprising a GENSET gene sequence, or portion thereof, in an antisense orientation relative to the direction of transcription, or by introducing antisense oligonucleotides directly into target cells. The therapeutic methodologies described herein are applicable to both human and non-human mammals.

5

Pharmaceutical and physiologically acceptable compositions

[0997] The present invention also relates to pharmaceutical or physiologically acceptable compositions comprising, as active agent, the polypeptides, nucleic acids or antibodies of the invention. The invention also relates to compositions comprising, as active agent, compounds selected 10 using the above-described screening protocols. Such compositions include the active agent in combination with a pharmaceutical or physiologically acceptable carriers such as a physiologically acceptable salt, ester, or salt of such esters. In the case of naked DNA, the "carrier" may be gold particles. The amount of active agent in the composition can vary with the agent, the patient and the effect sought. Likewise, the dosing regimen can vary depending on the composition and the 15 disease/disorder to be treated.

[0998] Therefore, the invention related to methods for the production of pharmaceutical composition comprising a method for selecting an active agent, compound, substance or molecule using any of the screening method described herein and furthermore mixing the identified active agent, compound, substance or molecule with a physiologically acceptable carrier.

[0999] The term "physiologically acceptable salts" refers to physiologically and pharmaceutically acceptable salts of the compounds of the invention: i.e., salts that retain the desired biological activity of the parent compound and do not impart undesired toxicological effects thereto. Physiologically acceptable base addition salts are formed with metals or amines, such as alkali and alkaline earth metals or organic amines. Examples of metals used as cations are sodium, potassium, 20 magnesium, calcium, and the like. Examples of suitable amines are N,N'-dibenzylethylenediamine, chloroprocaine, choline, diethanolamine, dicyclohexylamine, ethylenediamine, N-methylglucamine, and procaine (see, for example, Berge et al., "Pharmaceutical Salts," J. of Pharma Sci. (1977) 66:1-19). The base addition salts of said acidic compounds are prepared by contacting the free acid form with a sufficient amount of the desired base to produce the salt in the conventional manner. The free acid 25 form may be regenerated by contacting the salt form with an acid and isolating the free acid in the conventional manner. The free acid forms differ from their respective salt forms somewhat in certain physical properties such as solubility in polar solvents, but otherwise the salts are equivalent to their 30

respective free acid for purposes of the present invention. As used herein, a "pharmaceutical addition salt" includes a physiologically acceptable salt of an acid form of one of the components of the compositions of the invention. These include organic or inorganic acid salts of the amines. Preferred acid salts are the hydrochlorides, acetates, salicylates, nitrates and phosphates. Other suitable
5 physiologically acceptable salts are well known to those skilled in the art and include basic salts of a variety of inorganic and organic acids, such as, hydrochloric acid, hydrobromic acid, sulfuric acid or phosphoric acid; with organic carboxylic, sulfonic, sulfo or phospho acids or N-substituted sulfamic acids, for example acetic acid, propionic acid, glycolic acid, succinic acid, maleic acid, hydroxymaleic acid, methylmaleic acid, fumaric acid, malic acid, tartaric acid, lactic acid, oxalic acid, gluconic acid,
10 glucaric acid, glucuronic acid, citric acid, benzoic acid, cinnamic acid, mandelic acid, salicylic acid, 4-aminosalicylic acid, 2-phenoxybenzoic acid, 2-acetoxybenzoic acid, embonic acid, nicotinic acid or isonicotinic acid; and with amino acids, the 20 alpha-amino acids involved in the synthesis of proteins in nature, for example glutamic acid or aspartic acid, and also with phenylacetic acid, methanesulfonic acid, ethanesulfonic acid, 2-hydroxyethanesulfonic acid, ethane-1,2-disulfonic acid, benzenesulfonic
15 acid, 4-methylbenzenesulfonic acid, naphthalene-2-sulfonic acid, naphthalene-1,5-disulfonic acid, 2- or 3-phosphoglycerate, glucose-6-phosphate, N-cyclohexylsulfamic acid (with the formation of cyclamates), or with other acid organic compounds, such as ascorbic acid. Physiologically acceptable salts of compounds may also be prepared with a physiologically acceptable cation. Suitable physiologically acceptable cations are well known to those skilled in the art and include alkaline,
20 alkaline earth, ammonium and quaternary ammonium cations. Carbonates or hydrogen carbonates are also possible. For oligonucleotides, preferred examples of physiologically acceptable salts include but are not limited to (a) salts formed with cations such as sodium, potassium, ammonium, magnesium, calcium, polyamines such as spermine and spermidine, etc.; (b) acid addition salts formed with inorganic acids, for example hydrochloric acid, hydrobromic acid, sulfuric acid, phosphoric acid, nitric
25 acid and the like; (c) salts formed with organic acids such as, for example, acetic acid, oxalic acid, tartaric acid, succinic acid, maleic acid, fumaric acid, gluconic acid, citric acid, malic acid, ascorbic acid, benzoic acid, tannic acid, palmitic acid, alginic acid, polyglutamic acid, naphthalenesulfonic acid, methanesulfonic acid, p-toluenesulfonic acid, naphthalenedisulfonic acid, polygalacturonic acid, and the like; and (d) salts formed from elemental anions such as chlorine, bromine, and iodine.

30 [1000] The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to: parenteral, intracranial, intraorbital, intracapsular, intraspinal, intracisternal, intrapulmonary, oral, intravenous, intramuscular, intraarterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal,

enteral, topical, sublingual, rectal, intradermal, intravascular. In addition to the active ingredients, these pharmaceutical compositions may contain suitable physiologically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and 5 administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack PublishingCo. Easton, Pa).

[1001] Pharmaceutical compositions for oral administration can be formulated using physiologically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as powders, tablets, pills, 10 dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient. Suitable excipients are carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, or sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums including arabic and tragacanth; and proteins such as gelatin and collagen. If desired, disintegrating or solubilizing 15 agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a salt thereof, such as sodium alginate.

[1002] Formulations suitable for pulmonary or respiratory delivery include dry powders, liquid solutions or suspensions suitable for nebulization, and propellant formulations suitable for use in metered dose inhalers (MDI's). In typical embodiments, dry powder formulations will have a particle 20 size within a preferred range for deposition within the alveolar region of the lung, typically from $0.5\mu\text{m}$ to $5\mu\text{m}$. Respirable powders of pharmaceutical compositions within the preferred size range can be produced by a variety of conventional techniques, such as jet-milling, spray-drying, solvent precipitation, and the like. Dry powders can then be administered to the patient in conventional dry powder inhalers (DPI's) that use the patient's inspiratory breath through the device to disperse the 25 powder or in air-assisted devices that use an external power source to disperse the powder into an aerosol cloud (see, e.g., U.S. Pat. No. 5,458,135).

[1003] Liquid formulations for use in nebulizer systems preferably employ slightly acidic buffers (pH 4-6) such as acetate, ascorbate, and citrate, at concentrations of 5 mM to 50 mM. These buffers can act as antioxidants. Physiologically acceptable components to enhance or maintain 30 chemical stability include: antioxidants, chelating agents, protease inhibitors, isotonic modifiers, inert gases, and the like. A preferred type of nebulizer suitable for delivering such liquid formulations is described in U.S. Pat. No. 5,458,135.

[1004] For use in MDI's, the pharmaceutical composition will typically be processed into respirable particles as described for the dry powder formulations, and the particles then suspended in a suitable aerosol propellant (such as a CFC or HFC), typically being coated with a surfactant to enhance their dispersion. Such aerosol propellant formulations may further include a lower alcohol, such as 5 ethanol (up to 30% by weight) and other additives to maintain or enhance chemical stability and physiological acceptability (see, e.g., U.S. Patent 6,080,721).

[1005] For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

[1006] Pharmaceutical formulations suitable for parenteral administration may be formulated 10 in aqueous solutions, preferably in physiologically compatible buffers such as Hanks solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethylcellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or 15 synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions.

[1007] The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, 20 dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

[1008] In one embodiment, the preparation is a lyophilized powder which may contain any or all of the following: 1-50 mM histidine, 0.1%-2% sucrose, and 2-7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

[1009] After pharmaceutical compositions have been prepared, they can be placed in an 25 appropriate container and labeled for treatment of an indicated condition. For administration of a GENSET polypeptide, such labeling would include, e.g., amount, frequency, and method of administration.

[1010] Pharmaceutical compositions suitable for use in the invention include compositions 30 wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art. For any compound, a therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models, usually mice, rabbits, dogs, or pigs. The animal model may

also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

[1011] A therapeutically effective dose refers to that amount of active ingredient, for example a GENSET polypeptide or fragments thereof, antibodies specific to GENSET polypeptides, agonists, 5 antagonists or inhibitors of GENSET polypeptides, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., ED50 (the dose therapeutically effective in 50% of the population) and LD50 (the dose lethal to 50% of the population). The dose ratio between therapeutic and toxic effects is the therapeutic index, and it can be expressed as the ratio, LD50/ED50.
10 Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, sensitivity of the patient, and the route of administration.

15 [1012] The exact dosage will be determined by the practitioner, in light of factors related to the subject that requires treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, general health of the subject, age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and
20 tolerance/response to therapy. Other factors that may be considered when evaluating the proper dosage include the chemical nature of the compound destined for delivery, the biological responses associated with the compound (both intended and coincidental) and anticipated contraindications. Additionally, the mode of delivery, the duration and frequency of administration (e.g. n doses per hours, n doses per day, n doses per week, cumulative dosage per day, cumulative dosage per week),
25 the biologically effective dose delivered to target site, often indicated by plasma level concentrations, and the rate or efficiency of compound clearance from the body may be considered. Long-acting pharmaceutical compositions may be administered, e.g., every 3 to 4 days, every week, or once every two weeks depending on half-life and clearance rate of the particular formulation.

[1013] Normal dosage amounts may vary depending upon the route of administration.
30 Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc. In general, for a 75kg individual the

normal dosage range are as follows: for a small molecule compound an effective dose is usually between 0.3–50 mg/kg; for recombinant polypeptides an effective dose is usually between 0.25–7.5 mg/kg; for compounds used for mediating humoral immune responses (e.g., polyvalent pneumococcal vaccine, Rh_o (D) immune globulin, Hepatitis B vaccine, anti-CD20 antigen) the effective dose is 5 usually between 0.0015-1.5 mg/kg; for hormone supplemental compounds (e.g. estradiol, norethindrone) the effective dose is usually between 0.0005-0.5 mg/kg depending upon delivery system utilized (e.g. transdermal, oral, topical).

[1014] Transdermal delivery systems (e.g. estradiol transdermal system, transdermal scopolamine system, transmural nicotine patch) must be calibrated for nominal delivery dosages 10 based upon efficiency of percutaneous delivery for the individual and specific compounds, surface area (cm²) of transdermal system contact, quantity and form of compound integrated into transdermal delivery system and anatomical location of positioned transdermal system. The effective dosage range of compounds administered in this manner is usually between 0.005-0.5 mg/kg

15 Uses of Genset Sequences: Computer-Related Embodiments

[1015] It will be appreciated by those skilled in the art that the nucleic acid codes of the invention and polypeptide codes of the invention can be stored, recorded, and manipulated on any medium which can be read and accessed by a computer. As used herein, the words “recorded” and “stored” refer to a process for storing information on a computer medium. A skilled artisan can 20 readily adopt any of the presently known methods for recording information on a computer readable medium to generate manufactures comprising one or more of the nucleic acid codes of the invention, or one or more of the polypeptide codes of the invention. Another aspect of the present invention is a computer readable medium having recorded thereon at least 1, 2, 5, 10, 15, 20, 25, 30, or 50 nucleic acid or polypeptide codes of the invention.

25 [1016] Computer readable media include magnetically readable media, optically readable media, electronically readable media and magnetic/optical media. For example, the computer readable media may be a hard disk, a floppy disk, a magnetic tape, CD-ROM, Digital Versatile Disk (DVD), Random Access Memory (RAM), or Read Only Memory (ROM).

[1017] Embodiments of the present invention include systems, particularly computer systems 30 which store and manipulate the sequence information described herein. As used herein, “a computer system” refers to the hardware components, software components, and data storage components used to analyze the nucleotide sequences of the nucleic acid codes of the invention or the amino acid

sequences of the polypeptide codes of the invention. The computer system preferably includes a processor for processing, accessing and manipulating the sequence data. The processor can be any well-known type of central processing unit, such as the Pentium III from Intel Corporation, or similar processor from Sun, Motorola, Compaq or International Business Machines. Preferably, the computer 5 system is a general purpose system that comprises the processor and one or more internal data storage components for storing data, and one or more data retrieving devices for retrieving the data stored on the data storage components. A skilled artisan can readily appreciate that any one of the currently available computer systems are suitable.

[1018] In some embodiments, the computer system may comprise a sequence comparer for 10 comparing the above-described nucleic acid codes of the invention or the polypeptide codes of the invention stored on a computer readable medium to reference nucleotide or polypeptide sequences stored on a computer readable medium or present in a database. A "sequence comparer" refers to one or more programs which are implemented on the computer system to compare a nucleotide or polypeptide sequence with other nucleotide or polypeptide sequences and/or compounds including but 15 not limited to peptides, peptidomimetics, and chemicals stored within the data storage means. For example, the sequence comparer may compare the nucleotide sequences of nucleic acid codes of the invention or the amino acid sequences of the polypeptide codes of the invention stored on a computer readable medium to reference sequences stored on a computer readable medium to identify homologies, motifs implicated in biological function, or structural motifs. The various sequence 20 comparer programs identified elsewhere in this patent specification are particularly contemplated for use in this aspect of the invention. Alternatively, a nucleotide or amino acid sequence is compared with a database of sequences in order to determine the homology levels between the sequence and the sequences in the database. The database of sequences can be a private database stored within the computer system, or a public database such as GENBANK, PIR OR SWISSPROT that is available, 25 e.g., through the Internet.

[1019] Accordingly, one aspect of the present invention is a computer system comprising a processor, a data storage device having stored thereon a nucleic acid code of the invention or a polypeptide code of the invention, a data storage device having retrievably stored thereon reference nucleotide sequences or polypeptide sequences to be compared to the nucleic acid code of the 30 invention or polypeptide code of the invention and a sequence comparer for conducting the comparison.

[1020] Alternatively, the computer program may be a computer program which compares the nucleotide sequences of the nucleic acid codes of the present invention, to reference nucleotide

sequences in order to determine whether the nucleic acid code of the invention differs from a reference nucleic acid sequence at one or more positions. Optionally such a program records the length and identity of inserted, deleted or substituted nucleotides with respect to the sequence of either the reference polynucleotide or the nucleic acid code of the invention. In one embodiment, the computer
5 program may be a program which determines whether the nucleotide sequences of the nucleic acid codes of the invention contain one or more single nucleotide polymorphisms (SNP) with respect to a reference nucleotide sequence. These single nucleotide polymorphisms may each comprise a single base substitution, insertion, or deletion. The method may also be performed by reading at least 2, 5,
10, 15, 20, 25, 30, or 50 of the nucleic acid codes of the invention and the reference nucleotide sequences through the use of the computer program and identifying differences between the nucleic acid codes and the reference nucleotide sequences with the computer program.

[1021] In other embodiments the computer based system may further comprise an identifier for identifying features within the nucleotide sequences of the nucleic acid codes of the invention or the amino acid sequences of the polypeptide codes of the invention. An "identifier" refers to one or
15 more programs which identifies certain features within the above-described nucleotide sequences of the nucleic acid codes of the invention or the amino acid sequences of the polypeptide codes of the invention. In one embodiment, the identifier may comprise a program which identifies an open reading frame in the cDNAs codes of the invention.

[1022] In additional embodiments, the present amino acid sequences can be analyzed using
20 any of a number of computer programs to identify structural features of the protein, such as secondary, tertiary, and quarternary structures, to identify potential binding partners, etc. The results of the molecular modeling analysis may then be used, e.g., in rational drug design techniques to identify agents which modulate the activity of the polypeptide codes of the invention.

[1023] Accordingly, another aspect of the present invention is a method of identifying a
25 feature within the nucleic acid codes of the invention or the polypeptide codes of the invention comprising reading the nucleic acid code(s) or the polypeptide code(s) through the use of a computer program which identifies features therein and identifying features within the nucleic acid code(s) or polypeptide code(s) with the computer program. The method may be performed by reading a single sequence or at least 2, 5, 10, 15, 20, 25, 30, or 50 of the nucleic acid codes of the invention or the
30 polypeptide codes of the invention through the use of the computer program and identifying features within the nucleic acid codes or polypeptide codes with the computer program.

[1024] Motifs which may be detected using the above programs include leucine zippers, helix-turn-helix motifs, glycosylation sites, ubiquitination sites, alpha helices, beta sheets, signal sequences, sequences implicated in transcription regulation such as homeoboxes, acidic stretches, enzymatic active sites, substrate binding sites, and enzymatic cleavage sites.

5

[1025] Although this invention has been described in terms of certain preferred embodiments, other embodiments which will be apparent to those of ordinary skill in the art in view of the disclosure herein are also within the scope of this invention. Accordingly, the scope of the invention is intended to be defined only by reference to the appended claims.

10 [1026] Throughout this application, various publications, kit manuals, patents and published patent applications are cited. The entire disclosures of each of these publications, patents, manuals and published patent specifications are hereby incorporated by reference into the present disclosure to more fully describe the state of the art to which this invention pertains.

Table I

SEQ ID NO.	Sequence Type	Clone ID_Clone Name	Name	ATCC Deposit	ATCC Deposit Date
1	DNA	243525_116-119-1-0-G6-F	PCNAlt		
2	Protein	243525_116-119-1-0-G6-F	PCNAlt		
3	DNA	643144_181-17-2-0-A12-F	Lectir		
4	Protein	643144_181-17-2-0-A12-F	Lectir		
5	DNA	212950.cREC_116-075-2-0-H1-F	v1ADAM20		
6	Protein	212950.cREC_116-075-2-0-H1-F	v1ADAM20		
7	DNA	1000849866_181-44-3-0-A9-F	Lipoglobulin		
8	Protein	1000849866_181-44-3-0-A9-F	Lipoglobulin		
9	DNA	164341_117-001-5-0-E2-F	PBK		
10	Protein	164341_117-001-5-0-E2-F	PBK		
11	DNA	1000837037_228-43-2-0-C3-F	Myeloidin		
12	Protein	1000837037_228-43-2-0-C3-F	Myeloidin		
13	DNA	101005_105-020-4-0-H11-F	vACT		
14	Protein	101005_105-020-4-0-H11-F	vACT		
15	DNA	500743419_188-281-3-0-H5-F	Claudinyn-5		
16	Protein	500743419_188-281-3-0-H5-F	Claudinyn-5		
17	DNA	645730_181-16-1-0-G9-F	Benzodiazepine receptor 3 (BZRP-R3)		
18	Protein	645730_181-16-1-0-G9-F	Benzodiazepine receptor 3 (BZRP-R3)		
19	DNA	646762_181-21-2-0-A3-F	Benzodiazepine receptor 4 (BZRP-R4)		
20	Protein	646762_181-21-2-0-A3-F	Benzodiazepine receptor 4 (BZRP-R4)		
21	DNA	420594_145-19-4-0-E7-F	Scolakin		
22	Protein	420594_145-19-4-0-E7-F	Scolakin		
23	DNA	119658_105-067-2-0-H4-F	DOV		
24	Protein	119658_105-067-2-0-H4-F	DOV		

SEQ ID NO.	Sequence Type	Clone ID_Clone Name	Name	ATCC Deposit	ATCC Deposit Date
25	DNA	500706437_204-5-2-0-C9-F	Placentalin		
26	Protein	500706437_204-5-2-0-C9-F	Placentalin		
27	DNA	176355_117-005-2-0-H11-F	NAAR		
28	Protein	176355_117-005-2-0-H11-F	NAAR		
29	DNA	222588_116-094-1-0-H2-F	Neurexinal		
30	Protein	222588_116-094-1-0-H2-F	Neurexinal		
31	DNA	119033_105-066-4-0-F10-F	NPAISY		
32	Protein	119033_105-066-4-0-F10-F	NPAISY		
33	DNA	125402_105-074-4-0-F3-F	vCRTL-1		
34	Protein	125402_105-074-4-0-F3-F	vCRTL-1		
35	DNA	107640_105-036-2-0-H3	LIRION		
36	Protein	107640_105-036-2-0-H3	LIRION		
37	DNA	588394_160-105-40-A11-F	SLAMP		
38	Protein	588394_160-105-40-A11-F	SLAMP		
39	DNA	495718_160-26-2-0-E12-F	SAP-MU-10		
40	Protein	495718_160-26-2-0-E12-F	SAP-MU-10		
41	DNA	612386_187-9-4-0-B2-F	Cytogram		
42	Protein	612386_187-9-4-0-B2-F	Cytogram		
43	DNA	1000838982_220-20-4-0-B7-F	Tetranab		
44	Protein	1000838982_220-20-4-0-B7-F	Tetranab		
45	DNA	500720840_205-20-1-0-B7-F	PDI		
46	Protein	500720840_205-20-1-0-B7-F	PDI		
47	DNA	146821_106-020-1-0-G3-F	NBART		
48	Protein	146821_106-020-1-0-G3-F	NBART		
49	DNA	644724_181-21-1-0-A12-F	NBTG		
50	Protein	644724_181-21-1-0-A12-F	NBTG		
51	DNA	583702_181-8-4-0-C8-F	vITH1		
52	Protein	583702_181-8-4-0-C8-F	vITH1		

5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100

Table II

SEQ ID NO:	ORF	Signal Peptide	Mature peptide	Polyadenylation Signal	PolyA tail
1	[225-593]	-	-	[769-774]	[800-850]
3	[91-777]	-	-	[826-831]	[848-883]
5	[235-2562]	-	-	-	-
7	[264-926]	[264-320]	[321-926]	[1404-1409]	[1421-1436]
9	[93-551]	[93-197]	[198-551]	[2085-2090]	[2117-2132]
11	[201-986]	[201-269]	[270-986]	[2233-2238]	[2251-2266]
13	[66-1334]	[66-134]	[135-1334]	[1558-1563]	[1582-1597]
15	[153-806]	-	-	[1362-1367]	[1382-1397]
17	[63-572]	-	-	[750-755]	[767-782]
19	[63-572]	-	-	[750-755]	[774-789]
21	[86-403]	-	-	[504-509]	[540-555]
23	[415-1653]	-	-	-	[1688-1726]
25	[76-339]	[76-147]	[148-339]	[903-908]	[926-941]
27	[21-1118]	[21-89]	[90-1118]	[1858-1863]	[1879-1894]
29	[32-559]	-	-	[702-707]	[728-742]
31	[4-1533]	-	-	[1709-1714]	[1744-1766]
33	[11-802]	[11-64]	[65-802]	[836-841]	[862-877]
35	[38-1378]	[38-106]	[107-1378]	[107-1378]	-
37	[330-1478]	[330-398]	[399-1478]	[1722-1727]	[1742-1757]
39	[81-1517]	-	-	[2786-2791]	[2804-2818]
41	[121-546]	[121-231]	[232-546]	[739-744]	[755-770]
43	[136-501]	-	-	[1232-1237]	[1255-1340]
45	[118-1632]	[118-189]	[190-1632]	[1937-1942]	[1956-1999]
47	[154-546]	-	-	-	[722-836]
49	[196-708]	-	-	-	[847-862]
51	[62-778]	[62-124]	[125-778]	[912-917]	[932-937]

Table III

SEQ ID NO:	Positions of immunogenic epitopes
4	1..8:12..17:57..64:63..74:90..114:122..132:162..174:186..194:196..206:207..214
6	3..19:20..28:79..89:100..107:150..164:209..221:333..343:368..373:398..406:418..432: 430..439:447..453:513..521:524..539:553..565:569..583:582..593: 635..648:684..692:694..733:770..775
8	22..31:39..45:81..90:94..99:98..106:108..116:136..144:183..191: 190..196:205..217
10	11..21:50..60:77..84:118..126
12	81..88:142..150:180..186:194..203:219..227:251..261
14	25..31:34..45:124..133:182..193:216..221:219..230:230..239:267..272:279..285: 296..302:317..325:416..422
16	58..63:188..194:199..217
18	22..36:151..156:161..169
20	22..28:31..36:151..156:161..169
22	21..31:71..77:77..105
24	1..14:20..44:50..59:82..99:101..113:153..162:163..169:175..194:254..259:291..298: 326..331:353..361:368..378:381..386:394..406
26	24..33:52..62:67..78
28	21..42:79..91:122..132:182..189:264..277:286..293:314..321:326..336:354..363
30	1..17:22..28:30..42:48..55:93..101
32	28..35:63..75:72..82:87..95:152..162:176..185:196..212:224..243:262..270:355..363: 369..374:407..417:455..462:471..480:489..503
34	50..59:77..82:88..98:108..120:208..216:234..241
36	16..22:63..83:87..112:145..151:168..174:189..195:221..232:236..244:281..291: 290..307:304..320:334..341:344..355:361..373:418..433
38	29..39:39..44:51..59:71..79:99..108:111..120:139..145:152..157:164..173:185..195: 217..232:244..257:264..270:277..286:331..341:352..365:371..381
40	24..29:34..41:52..58:92..104:124..131:180..198:332..340:403..411:422..433:446..455
42	66..73:91..102:113..130
46	31..41:38..43:58..63:86..98:107..122:128..133:176..186:193..202:209..215:266..283: 341..349:354..368:368..378:390..397:406..419:421..427:440..447:456..471:480..485

SEQ ID NO:	Positions of immunogenic epitopes
48	98..105:123..130
50	30..40:53..68:114..119:128..139:157..166
52	15..31:46..58:107..113:150..155:172..184:197..210